DOCUMENT RESUME

ED 032 812 48 FL 001 480

By-Wingard, Harold B.

A Study of Audiolingual Instruction Needed by Sixth-Grade Pupils Prior to Introducing Reading in First Year Spanish. Final Report.

San Diego Unified School District, Calif.

Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.

Bureau No-BR-6-1362

Pub Date Aug 68

Contract-OEC-4-7-061362-2709

Note-90p.

EDRS Price MF-$0.50 HC-$4.60

Descriptors-*Audiolingual Methods, Audiolingual Skills, Course Content, Elementary School Students, *Fles Programs, Grade 6, Language Instruction, *Language Research, Language Tests, *Reading Instruction, Reading Skills, Research Needs, Research Projects, *Spanish, Tables (Data), Teaching Methods

Identifiers-San Diego Unified School District

This cooperative research project, supported by the Office of Education Bureau of Research, attempts to answer questions on the effectiveness of reading instruction during a sixth grade beginning Spanish program. Major areas studied include: (1) the effects of reading instruction on listening, speaking, and reading skills, (2) the amounts of reading instruction necessary to effect a significant difference, and (3) the relative effectiveness of new and review reading content. Statistically inconclusive results limit the conclusions to a discussion of trends. The major recommendation growing out of this study is that further investigation in all three areas is greatly needed. Extensive appendixes are included. (AF)

ED0 32812

BR 6-1362
PA 48
OE/BR

FINAL REPORT
Cooperative Research Project
Contract No. OEC-4-7-061362-2709

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A STUDY OF AUDIOLINGUAL INSTRUCTION NEEDED BY SIXTH-GRADE PUPILS PRIOR TO INTRODUCING READING IN FIRST YEAR SPANISH

August 1968

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

FL 001 480

FINAL REPORT
Cooperative Research Project
Contract No. OEC-4-7-061362-2709

SAN DIEGO UNIFIED SCHOOL DISTRICT

Dr. Jack Hornback
Superintendent

Dr. William H. Stegeman
Assistant Superintendent
Curriculum Services Division

Principal Investigator	Harold B. Wingard Foreign Language Specialist
Project Coordinator	Grace W. Perkins Principal Lowell Elementary School
Project Research Coordinator	William H. Vogler Director of Testing Services
Project Research Consultants	Dr. William J. Moonan Director of Program Development National Computer Systems
	Dr. Walter V. Romanowski Vice Principal Bell Junior High School

A STUDY OF AUDIOLINGUAL INSTRUCTION NEEDED BY SIXTH-GRADE PUPILS PRIOR TO INTRODUCING READING IN FIRST YEAR SPANISH

Cooperative Research Project
Contract No. OEC-4-7-061362-2709

August 1968

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

San Diego Unified School District

San Diego, California

TABLE OF CONTENTS

Page

LIST OF TABLES AND FIGURES iv

I INTRODUCTION 1

A. Nature of the Problem 1
B. Objectives and Null Hypotheses of the Study 2
C. Limitations of the Study 3

II METHODS AND PROCEDURES 4

A. The Experimental Design 4
B. Selection and Sampling Procedures and Variables 7
C. Instructional Materials, In-Service Training and Observation of Classes 13
D. The Experimental Model and Its Analysis 18
E. Test Construction and Testing Procedures 22
F. Chronological Overview 35

III EXPERIMENTAL RESULTS 41

A. Listening Test Results 41
B. Reading Test Results 43
C. Speaking Test Results 46

IV DISCUSSION OF THE RESULTS 49

A. Hypotheses 49
B. Methods Differences 49

V CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS 52

A. Conclusions 52
B. Implications 52
C. Recommendations 53

VI SUMMARY 55

APPENDIX A A-1

APPENDIX B B-1

APPENDIX C C-1

APPENDIX D D-1

LIST OF TABLES AND FIGURES

Table No. Page

I Experimental Design 5

II Teacher-Class Assignment to Groups by Stratified Random Selection 11

III Types of Tests Administered at Each Testing Period . . . 23

IV Speaking Tests 27

V Number of Discrete Phonological Items per Test 35

VI Means for Listening Scores 42

VII ANACOVA Table for Listening Results 42

VIII Means for Reading Scores 43

IX ANACOVA Table for Reading Results 44

X Multiple Comparison on Reading Test Results 45

XI Raw Score Data for Speaking Scores 46

XII Means for Speaking Scores 47

XIII ANACOVA Table for Speaking Results 47

XIV Comparison of Methods 48

XV Comparison of Raw Score Differences 50

Figure No.

1 Equipment and Seating Arrangement for Individual Testing 29

CHAPTER I

INTRODUCTION

A. Nature of the Problem

The California State Legislature in 1961 enacted into law a bill which required that foreign language instruction be given to all children in grades 6, 7, and 8. Although the legislation was not to take effect until the 1965-66 school year, interest in foreign language instruction at the elementary school level was greatly stimulated by the new enactment.

Many school districts within the state began to initiate programs. Numerous questions concerning materials and techniques for instruction were raised.

In the San Diego Unified School District, Spanish instruction for sixth-grade pupils was initiated district-wide in the 1965-66 school year. The district adopted and modified Primer Curso (Holt, Rinehart and Winston: New York, 1964). Although the publishers incorporated reading instruction, the district program followed prior to this study was strictly an audiolingual one.

Since the initiation of Spanish instruction in the district, teachers have questioned the omission of reading in the program. Their interest in introducing reading and using the textbook has been strong. Among their reasons are the following:

1. "The printed word reinforces the pupil's listening and speaking skills."
2. "The pupils are ready to read; to deny instruction in this realm is to deny development of the pupil's potential."
3. "Sixth-grade pupils are accustomed to using texts; subjects in which texts are not used are not considered important by the pupils at this level."
4. "Pupils ask to see how the utterances look in writing."
5. "If the printed word is withheld, pupils write their concept of the word on their own initiative. In that the utterances are often misspelled by the pupil, incorrect learnings are acquired."
6. "Pupils tend to become bored with purely audiolingual instruction; the addition of reading would add variety and enrichment to the program.

B. Objectives and Null Hypotheses of the Study

In response to teacher and pupil interest in starting reading instruction in the sixth-grade Spanish program, the San Diego Unified School District undertook an 18-month project concerning this issue. The study, initiated in March, 1967, was designed to answer two primary questions:

1. Does the inclusion of reading in a program of sixth-grade beginning Spanish instruction affect achievement in the skills of listening, speaking, and reading?

2. If the inclusion of reading results in a significant difference in achievement, what amount of reading instruction results in the most significant difference?

The study was also designed to answer a secondary question:

> Does a significant difference in achievement take place when reading instruction utilizes a combination of new and review content rather than content which is entirely review?

While all of the above research hypotheses were tested in this project, the main objective to which the project addressed itself was the determination of what combination of prereading instruction and instruction involving reading in sixth-grade beginning Spanish classes results in maximum achievement in listening, speaking, and reading, when groups of students, each studying a specified amount of content for a specified amount of time according to a specified learning procedure, receive varying amounts of prereading instruction. The following null hypotheses were tested:

1. Major Hypothesis:

   ...There will be no differences in listening, reading, and speaking test achievement among the various groups to be studied in the project.

2. Secondary Hypotheses:

   ...There will be no differences in listening and reading test achievement dependent upon the ability of the teacher. (An attempt was made to make the effect of the program materials independent of teacher ability.)

   ...There will be no difference in listening and reading test achievement between boys and girls.

C. Limitations of the Study

One of the weaknesses in this study was in the materials themselves. The sequence of content was such that the various grammatical structures introduced throughout the course of instruction were essentially at the same level of difficulty. Furthermore, the structures at later stages were not dependent upon those from earlier stages. Thus, it is likely that the study would have revealed the same results had it not been conducted for as extensive period as 27 weeks.

The remainder of this report will consider methods and procedures, experimental results, and general conclusions.

CHAPTER II

METHODS AND PROCEDURES

A. The Experimental Design

In the planning phase of the project, two modifications of the experimental design were made. These modifications served to tighten the controls on the major variables and provide controls on other variables of significance. The description below is of the modified design which was actually used.

Nine groups, each composed of 8 classes, participated in the study. Each group received daily instruction in Spanish, varying from 15 to 20 minutes. The duration of the data collection period was for the 27 instructional weeks from October 23, 1967 to June 6, 1968.

The variations in treatment are described below and in Table I.

1. Method 1: Oral instruction 15 minutes daily, reading instruction 5 minutes daily for 27 weeks.

   Reading instruction on the new material of each lesson was given from the first day of instruction until the end of the experiment. Reading review of previously introduced material was also presented.

2. Method 2: Oral instruction 15 minutes daily for 27 weeks.

   This group followed the same oral program as Method 1. No instruction was given for the 5-minute period in which Method 1 had reading instruction. Data from Method 2 provided for more definite isolation of the effect of reading instruction in Method 1.

3. Method 3: Oral instruction 20 minutes daily for the first 9 weeks.

   Oral instruction 15 minutes daily, reading instruction 5 minutes daily for the remaining 18 weeks.

   This group began reading after 9 weeks of oral instruction. Reading instruction was given on the new material of each lesson. Reading review of previously introduced material was also presented.

4. Method 4: Oral instruction 20 minutes daily for the first 9 weeks.

   Oral instruction 15 minutes daily for the remaining 18 weeks.

COOPERATIVE RESEARCH PROJECT

TABLE I

Experimental Design

0 weeks 9 weeks 18 weeks 27 weeks

<table>
<tr><th>METHODS</th><th></th><th></th><th></th></tr>
<tr><td rowspan="2">1</td><td colspan="3">Oral instruction 15 minutes per day</td></tr>
<tr><td colspan="3">Reading instruction 5 minutes per day</td></tr>
<tr><td rowspan="2">2</td><td colspan="3">Oral instruction 15 minutes per day</td></tr>
<tr><td colspan="3">No instruction 5 minutes per day</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">Oral instruction 20 minutes per day</td><td colspan="2">Oral instruction 15 minutes per day</td></tr>
<tr><td colspan="2">Reading instruction current material 5 minutes per day</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2">Oral instruction 20 minutes per day</td><td colspan="2">Oral instruction 15 minutes per day</td></tr>
<tr><td colspan="2">No instruction 5 minutes per day</td></tr>
<tr><td rowspan="2">5</td><td rowspan="2">Oral instruction 20 minutes per day</td><td colspan="2">Oral instruction 15 minutes per day</td></tr>
<tr><td colspan="2">Reading instruction review material 5 minutes per day</td></tr>
<tr><td rowspan="2">6</td><td colspan="2" rowspan="2">Oral instruction 20 minutes per day</td><td>Oral instr. 15 min. per day</td></tr>
<tr><td>Rdg.inst.cur.mat. 5 min/day</td></tr>
<tr><td rowspan="2">7</td><td colspan="2" rowspan="2">Oral instruction 20 minutes per day</td><td>Oral instr. 15 min. per day</td></tr>
<tr><td>No instr. 5 min. per day</td></tr>
<tr><td rowspan="2">8</td><td colspan="2" rowspan="2">Oral instruction 20 minutes per day</td><td>Oral instr. 15 min. per day</td></tr>
<tr><td>Rdg.inst.rev.mat. 5 min/day</td></tr>
<tr><td>9</td><td colspan="3">Oral instruction 20 minutes per day</td></tr>
</table>

* Nine methods, 8 classes in each group.

This group followed the same oral program as Methods 3 and 5. No instruction was given in the 5-minute period in which Methods 3 and 5 had reading instruction. Data from Method 4 provided for more definite isolation of the effect of the reading instruction in Methods 3 and 5.

5. **Method 5**: Oral instruction 20 minutes daily for the first 9 weeks.

   Oral instruction 15 minutes daily, reading instruction 5 minutes daily for the remaining 18 weeks.

   This group began reading after 9 weeks of oral instruction. Reading instruction was given on a special block of previously introduced material.

6. **Method 6**: Oral instruction 20 minutes daily for the first 18 weeks.

   Oral instruction 15 minutes daily, reading instruction 5 minutes daily for the remaining 9 weeks.

   This group began reading after 18 weeks of oral instruction. Reading instruction was given on the new material of each lesson. Reading review of previously introduced material was also presented.

7. **Method 7**: Oral instruction 20 minutes daily for the first 18 weeks.

   Oral instruction 15 minutes daily for the remaining 9 weeks.

   This group followed the same oral program as Methods 6 and 8. No instruction was given in the 5-minute period in which Methods 6 and 8 had reading instruction. Data from Method 7 provided for more definite isolation of the effect of the reading instruction in Methods 6 and 8.

8. **Method 8**: Oral instruction 20 minutes daily for the first 18 weeks.

   Oral instruction 15 minutes daily, reading instruction 5 minutes daily for the remaining 9 weeks.

   This group began reading after 18 weeks of oral instruction. Reading instruction was given on a special block of previously introduced material.

9. **Method 9**: Oral instruction 20 minutes daily for 27 weeks.

This group was the control method for the experiment and followed the current San Diego City Schools Spanish program for classes not participating in the project.

Thus, in this design the following factors were held constant in all methods:

1. The amount of content covered. In Methods 2, 4, and 7, content was not as intensively practiced as in the other groups.

2. The total amount of time devoted to Spanish instruction, except in Methods 2, 4, and 7. Method 2 was created as a special control for Method 1; Method 4 for Methods 3 and 5; Method 7 for Methods 6 and 8.

The following factors were varied:

1. Primary factors

   a. The length of time devoted to oral instruction.

   b. The length of time devoted to reading instruction.

2. Secondary factor: The content used for reading instruction.

   In Methods 1, 3, and 6, the reading content was the same as the oral content of the lesson. It consisted of the new material of the lesson and a block of review material systematically selected from previously taught lessons. In Methods 5 and 8, the oral content was the same as the oral content of Methods 1, 3, and 6. However, the reading content was entirely different. It consisted of the new and review reading material of the lesson introduced 35 days earlier. For example, the Methods 5 classes studying Unit I, Lesson 36 were given reading instruction on the content of Unit I, Lesson 1.

B. Selection and Sampling Procedures and Variables

Selection of Teachers

Originally, the selection of teachers was to be based upon their successful qualification as Spanish teachers. This qualification was to be predicated upon scores obtained on the MLA Proficiency Test for Foreign Language and upon other criteria established by the Modern Language Association. However, this plan proved untenable, in view of the limitations of time, project staffing, teacher competency and teacher attitude. The only workable method, under the circumstances, was to invite participation and make selection from the group which volunteered.

A letter was sent to all elementary school principals in the district, inviting them and their Spanish teachers to attend a regional orientation meeting concerning the project. Four orientation meetings were conducted by the staff. In these meetings, the project was explained in detail, and teachers were invited to participate.

The group from which selection was made was comprised of 88 teachers. 82 teachers volunteered at the orientation meetings held late in the 1966-67 school year, and their continued interest was confirmed by their principals at the beginning of the 1967-68 school year. In addition, 6 teachers contacted the project office in the fall and asked to be included in the group. The 88 teachers were assigned to 60 elementary schools in the district. Prior to the final selection, a letter requesting confirmation of willingness to participate was sent to all principals whose staff members had volunteered.

Data concerning the racial and ethnic distribution of enrollment by school was obtained. Two schools were eliminated from the study on the basis of their high percentage of Mexican-American population (as indicated by surname). These percentages were 87.5% and 49.7%. This operation reduced the volunteer group to 85 teachers assigned to 58 schools.

Selection of Classes

In order to obtain a representative sample of the learners, a stratified random sampling procedure was followed. Each method by this procedure contained approximately the same range of ability among its groups as any other method. This is shown in Appendix A.

The 58 schools to which the teachers were assigned were ranked by IQ as indicated by sixth-grade pupil achievement on the Lorge-Thorndike test administered in September, 1966. Ranked from high IQ to low IQ by school, each teacher was assigned a number from 01 to 85.

The teachers of gifted classes were isolated from the list and treated separately. These teachers were assigned, one to each method, through random selection. One teacher of the gifted was removed from the list because the class involved was a gifted fifth-sixth combination grade, whereas all other gifted classes concerned were sixth-grade.

All teachers of non-gifted classes were scheduled to teach at least one regular sixth-grade class or one regular fifth-sixth combination grade class during the 1967-68 school year. These teachers and their classes were stratified into the following groups:

1. School IQ per Lorge-Thorndike: 108-115. Teachers #01-#17. (Teachers 01, 04, 05, 07, 12, and 14 were removed for placement in the gifted class stratum.)

2. School IQ per Lorge-Thorndike: 106-107. Teachers #18-#34. (Teachers 29, 32, and 34 were removed for placement in the gifted class stratum.)

3. School IQ per Lorge-Thorndike: 103-105. Teachers #35-#45. (Teachers 35 and 42 were removed for placement in the gifted class stratum.)
4. School IQ per Lorge-Thorndike: 101-102. Teachers #46-#56.
5. School IQ per Lorge-Thorndike: 100. Teachers #57-#66.
6. School IQ per Lorge-Thorndike: 96-99. Teachers #67-#75.
7. School IQ per Lorge-Thorndike: 88-95. Teachers #76-#85.

Through random selection, one teacher from each IQ stratum was assigned to each of the nine methods.

The methods were then balanced according to type of class. Each method was to involve:

1. Six regular sixth-grade classes
2. One gifted sixth-grade class
3. One regular fifth-sixth combination grade class.

This balance was achieved through one exchange in assignment. However, at mid-year, this balance was disturbed when one class in Method 6 changed from a sixth-grade to a fifth-sixth combination, because of an enrollment shift.

When the initial assignment was completed, 72 teachers and 72 classes in 53 schools were to participate. Letters were sent to all 88 volunteer teachers advising them of their status as participants or non-participants in the study.

Certain changes in the initial selection of participating teachers and classes occurred which were not within the control of the staff. Prior to the start of the experimental period, four teachers withdrew for various reasons. Replacement in three cases was made through random selection from the group unassigned in the initial selection. In one case, there was no unassigned teacher available in the appropriate IQ stratum. However, a teacher in that stratum, who was already assigned and was to teach Spanish to more than one class, was willing to undertake a dual assignment in the project.

In addition, three teachers volunteered to serve as alternates. One of these, who was teaching more than one class, served as an alternate in one group and as a regular participant in another group.

After the start of the experimental period, one teacher initiated reading 5 weeks ahead of schedule. The class was replaced, but because of a lack of alternates in that IQ stratum, the balance in IQ stratification was disturbed.

As seen in Table II, upon completion of these adjustments and changes, 72 teachers and 74 classes in 56 schools were participating in the study.

Many teachers in the district give Spanish instruction to classes other than their own. Therefore, an attempt was made to balance the group composition so that each method would contain:

1. Five classes taught Spanish by their regular classroom teacher, and

2. Three classes taught Spanish by a teacher other than their regular teacher.

In this attempt, letters were sent to all project teachers specifying the class to be taught in the experiment. However, through misunderstanding, at least 6 of the 27 teachers who were to use a class other than their regular class, used their own class in the experiment.

Subsampling for Speaking Skills

Whereas the listening comprehension and reading comprehension tests were group tests administered class-wide to all methods, the batteries of speaking tests had to be administered individually to pupils. The length of the tests and the necessity for individual administration made it virtually impossible in the time available to give the tests to every pupil in the sample. Therefore, it was decided to select three pupils in each class to receive the speaking tests.

Three speaking test batteries were given during the data collection period, and the pupil participants were held constant whenever absenteeism and transfer did not preclude it.

The selection of pupils was made in the following way:

1. The classes were listed in alphabetical order by school name, and divided into two groups.

2. The teachers of the first group were instructed to test two boys and one girl; those of the second were to test two girls and one boy.

3. The selection was made from class lists. Teachers were asked to select pupils occupying a specific location on the class list. From the standpoint of validity, it was more desirable to test pupils who were enrolled from the outset of the program, so selection was made from the first decade of names on the class list.

TEACHER-CLASS ASSIGNMENT TO GROUPS BY STRATIFIED RANDOM SELECTION

Table II

Explanation of entries:
Line 1 - Teacher name
Line 2 - Teacher selection number and class type
Line 3 - School name

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Method 9
Gifted classes	Huffman 29 gftd. 6 Gage	Lytwyn 07 gftd. 6 Farnum	Powell 14 gftd. 6 Hearst	Torio 01 gftd. 6 Silver Gate	Severson 34 gftd. 6 Marvin	Vanian 04 gftd. 6 Jackson	Cassidy ** 91 gftd. 6 MacDowell	Paderewski 35 gftd. 6 Ross	Epler 32 gftd. 6 Marvin
IQ 108-115* Teachers 01-17 (minus gifted)	Pentis 11 reg. 6 Jones	Stallo 13 reg. 6 Jones	Webb 08 reg. 6 Farnum	Patterson 03 reg. 6 Jackson	Aragon 09 5/6 Longfellow	Messick 06 reg. 6 Bird Rock	von Borstel 16 reg. 6 Weinberger	Mankins 17 reg. 6 Lafayette	Atkins 15 reg. 6 Weinberger
IQ 106-107* Teachers 18-34 (minus gifted)	Parsley 24 reg. 6 Foster	Buckley 18 reg. 6 Riley	Hayes 21 reg. 6 Lindbergh Ross 27 reg. 6 Sessions	Koons 22 reg. 6 Juarez	Haley 19 reg. 6 Riley	Bonnell 26 reg. 6 Sessions	Lopez ** 31 reg. 6 Pacific Beach	Cornell 30- 5/6 Toler	Gray 28 reg. 6 Gage
IQ 103-105* Teachers 35-45 (minus gifted)	Shearin 37 reg. 6 Oak Park	Rondeau 36 5/6 Ross		Bandh 38 reg. 6 Cadman	Cassidy ** 42 reg. 6 MacDowell	Lutz 41 5/6 Barnard	Seymour-Scott 39 reg. 6 Andersen	Clingan 45 - 6 Paradise Hills	McKasson 43 reg. 6 Whitman
IQ 101-102* Teachers 46-56	McKay 53 reg. 6 Stevenson	Gills 51 reg. 6 Alcott	Soper 55-5/6 Birney	Schwartz 47 reg. 6 Bay Park	Fadem 54 reg. 6 Birney	Fernandez 46-5/6 Bay Park	Conner 49 reg. 6 Bv. Terr.	Dunlap 56 reg. 6 Webster	Astrab 48 -5/6 Bv. Terr.
IQ 100* Teachers 57-66	Osher 58 - 5/6 Hamilton	Fletcher 61 reg. 6 Hawthorne	Abbott 57 reg. 6 Hamilton	Pecoraro 66 -5/6 Darnall	McCray 60 reg. 6 Perry	Pena 64 reg. 6 Lee	Aceves 62 reg. 6 Brooklyn	Thompson 65 reg. 6 Keiller	Barkhurst 59 reg. 6 Perry
IQ 96-99* Teachers 67-75	Crockett 69 reg. 6 Euclid	D'amico 75 reg. 6 Carson	Purinton 74 reg. 6 Edison	Briggs 67 reg. 6 Audubon	Hopkins 70 reg. 6 Knox	Bond 71 reg. 6 Freese	Frye 73-5/6 Edison	Keefer 72 reg. 6 Linda V.	Lucas 68 reg. 6 Sequoia
IQ 88-95* Teachers 76-85	Parker 80 reg. 6 Dewey	Woods 78 reg. 6 Central	Haring 85 reg. 6 Sherman	Taylor 84 reg. 6 Balboa	Wright 79 reg. 6 Central	Druet 76 reg. 6 Fulton	Whitlock 82 reg. 6 Baker	Larson 77 reg. 6 Wshgton.	Sienko 81 reg. 6 Horton
Alternate	Lindgreen 25 reg. 6 Forward		~~27 reg. 6~~			Lopez ** 92 reg.6 Pac. Bch.			

* IQ of school as determined by 6th grade pupil achievement on Lorge-Thorndike test, administered fall 1966.
** Teacher is teaching two classes in experiment.

Variables

The dependent variables used as criteria were listening, reading and speaking tests administered at the close of the 27-week period. Analysis of covariance was used for the analysis. Independent covariates used as controls for each criterion variable were listening, reading, and speaking pretests, respectively, at the 9-week period; teacher ability ratings and scholastic ability of pupils were controls in each comparison. The experimental design was partly hierarchial and partly factorial in nature, involving sex, class groups, and methods.

In addition to the 9-to-27-week analysis, 9-to-18-, and 18-to-27-week analyses were planned. Initial inspection of 9-to-18-, and 18-to-27-week differences indicated that these comparisons would not have made any significant change in the findings of the overall 18-to-27-week analysis, with a saving of the time involved in comparing 864 group means.

Various measures were considered and initiated for purposes of classifying teacher proficiency. At the beginning of the project, the staff held frequent meetings in an effort to establish criteria for any subjective classification to be done. They also reviewed and considered use of the MLA-LA Spanish test and the Pimsleur Foreign Language Proficiency test as a measure of teacher ability.

Classification of the project teachers on the basis of scores obtained on such a standardized foreign language test was rejected, because such scores would not be a measure indicative of teacher competency in conducting the project program. The programmed instruction followed in the project was designed to be "teacher-proof".

At one time, the staff decided not to attempt any formal classification based upon observation and/or interviews. However, the Director of Testing Services for the district felt that first-hand data based upon teacher observation and/or interview was important for purposes of data interpretation.

For teacher ability a classification was based upon the following factors:

a. Mastery of subject

b. Lesson preparation

c. Rapport with pupils

d. Motivation and participation of pupils

On a composite of the four factors, each teacher was given one of the following ratings:

1. Outstanding

2. Average

3. Below Average

In some cases, teachers deviated from the prescribed program either through presenting material from the programs of other groups or through presenting material not incorporated in the curriculum. Variance such as this was indicated on the classification form.

C. Instructional Materials, In-Service Training and Observation of Classes

Instructional Materials

The following instructional materials were used by participating teachers:

Guide for Teaching Spanish, Grade Six,

Audio tapes accompanying the guide,

Primer Curso text books (Methods 1, 3, 5, 6, and 8),

Set of 121 visuals used in lieu of the publishers' charts.

In the original planning for the project, the district modification of Primer Curso (Holt, Rinehart and Winston: New York 1964) was to serve as the course of study.

No revision of the instructional material was anticipated; only an adjunct to the district program was to be written--a preliminary unit of nine lessons incorporating reading instruction.

A necessary step in preparation for the experiment was a close scrutiny of the instructional program. The staff, upon familiarizing themselves with content and programming, felt that a more structured programming was necessary for the conduct of such an experiment.

The partially written preliminary unit and accompanying tape were abandoned, and a series of staff meetings were held to plan the revision. It was decided by the Foreign Language Specialist of the District and the Curriculum Coordinator, Foreign Language/ESL of the San Diego County Schools, that the revision was to affect three facets of the program:

1. The course content was to be modified to provide a more structured sequence.

2. Sentence patterns and concepts which had proven excessively difficult for sixth-grade pupils were to be simplified.

3. The publishers' manual and the district guide were to be combined into one publication in order to simplify the teacher's task in preparing lessons.

The revised program was incorporated in the new curriculum guide, Guide for Teaching Spanish, Grade Six. The enormity of the task precluded its completion prior to the start of the 1967-68 school year. Writing of the guide was initiated in August, 1967 and, until May, 1968, the staff directed a major part of its efforts to this undertaking. This guide was published and distributed to teachers in sections. The exigencies of the project necessitated the hiring, at district expense, of extra personnel for the preparation of portions of the guide.

Staff meetings were held periodically to determine content of each unit of the curriculum. Content formulae, guidelines and fill-in forms for lesson plan writing, visual identification lists, translation sheets and proofreading sheets were developed by the staff to expedite the writing task. Also, in conjunction with the guide preparation, comprehensive itemizations of the content of the lessons were prepared. The staff also isolated a high percentage of the entities taught which would lend themselves to recombination into couplets or dialogue. Couplet and dialogue lists were then prepared as a source for more effective review segments in the lessons.

The total guide, issued in 7 sections, consisted of 153 lessons. The program, consisting of 5 units, was based upon the first 4 units of Primer Curso.

Of vital importance to the instructional program are the accompanying audiotapes. The sixth-grade teachers in the district generally possess minimum skills in the Spanish language. The success of their instruction is dependent, in many cases, upon the use of tapes to model the language.

Early in September, the staff began preparation of the tape scripts. The scripts were written in units throughout the year, as content was determined for sections of the guide.

The audiotapes, which covered the 153 lessons, were prepared and issued on 36 reels. On these tapes were presented the new material, the review material, and the special block of review material for Methods 5 and 8. A fixed number of utterance repetitions and of spaces for pupil repetition of the utterances was programmed on the audiotapes.

In order to employ native speakers of Spanish with excellent accent for recording on the tapes, the staff taped voice samples of 15 native speakers known to them. Among these were elementary and secondary teachers in the district, project personnel and members of the community. Nine adults were selected; of these, 5 participated routinely, and 3 were employed where additional personnel were needed for certain dialogues.

To add realism and interest to the tapes, Spanish-speaking children participated in the dialogues involving children. Early in the 1967-68 school year, two teachers of English as a Second Language began the screening for these pupil participants. From their classes with a large Mexican-American population, they selected eight pupils to try out for the roles. Project personnel went to the schools and recorded the voices of the eight children. After evaluation by the staff, four voices were selected, and permission to participate was obtained from the parents of the children and the principals of their schools of attendance.

Taping sessions were held throughout the year as the tape scripts were readied for use. Approximately 29 hours were devoted to the children's recording, and 90 hours to adult recording. After the master tape was edited, tapes were dubbed and distributed to the schools.

As work in this realm progressed, certain changes and improvements were effected.

Initially the number of repetitions of the utterances varied. In the final three units of the tapes, a repetition formula was applied to each utterance, on the basis of its being new or review material, and its treatment in the lesson. Fill-in forms for script writing were developed which materially reduced the time required for this undertaking.

An analysis of the tape recorder needs of the project teachers was made; and the district loaned each project teacher a recorder for her own use, if she did not already have one.

2,200 copies of the Primer Curso textbooks were purchased for project use, and distributed to the appropriate groups.

Each project teacher was supplied a set of 121 district-prepared visuals. These replaced the publishers' charts, and were provided by the district for all Spanish teachers conducting the regular program as well.

In-Service Training

In the teachers' meetings held late in the 1966-67 school year, there was no attempt to initiate in-service training for the project. The purpose of these meetings was solely to present an overview of the project, generate teacher interest, and answer questions. In-service training was delayed until the fall to allow for adequate preparation by the project staff.

Three general orientation meetings were held in early fall, prior to starting Spanish instruction for the year. In these meetings, to which all sixth-grade Spanish teachers were invited, revisions in the program were explained and the new materials were introduced.

Following each of these general meetings, a special session was conducted for the project teachers. At that time, the overview of the project was explained in detail, the use of curriculum materials for each group was outlined, and the calendar for teaching, testing and in-service training was reviewed. This information, with a list of the teachers participating and their method assignments, was also compiled in a special packet for each teacher.

Make-up sessions were held for these meetings and for all in-service meetings conducted by the project, so that all participating teachers would have the same training.

Three series of in-service meetings for specific methods were held during the year. The in-service training for each method was scheduled approximately one week prior to the initiation of reading instruction in that group. Teachers of Methods 2, 4, 7, and 9 also attended, although reading was not a part of their instructional program. Methods 1, 2, and 9 received training in the first series; Methods 3, 4, and 5 in the second series, and Methods 6, 7, and 8 in the third series.

In the in-service meetings, the project staff described the relationship between the three methods in attendance. They reiterated the necessity of adhering to the specific program of the method. In that teachers had not had experience in the teaching of reading in the Spanish program, the following techniques were presented for introducing the printed word:

1. Presentation of the model of the utterance by audiotape immediately prior to the display of the printed utterance.

2. Removal of the printed utterance upon pupil mispronunciation.

3. Replay of the taped model and reintroduction of the printed material.

Detailed descriptions of the instructional program for each method were written and distributed. A pacing chart, to help assure adherence to the prescribed schedule of lesson presentation, was also distributed to all project teachers.

To minimize error in test administration, three identical meetings concerning pupil testing were held. The test formats and purposes were explained; the individual tests were previewed; and demonstrations of test administration by tape were presented.

Also, a series of three demonstration lessons for each area in the district were held during the 1967-68 school year. Both project teachers and teachers conducting the regular program were invited to attend. The purpose of these lessons was to demonstrate effective teaching techniques and appropriate uses of curriculum materials. The project staff interpreted the demonstrations to the guests and answered questions concerning the program.

In addition, the project staff conducted individual demonstration lessons for 52 teachers in the district. Of this group, 31 were project teachers. The staff also conferred with these teachers in an effort to answer any questions or help with any problems concerning the program.

Observation of Classes

The observation of project classes served two purposes. The original purpose of the observations was to determine:

1. That the program appropriate for the group was being taught,
2. That the instructional tapes were being used correctly where indicated, and
3. That the pacing of the curriculum was on schedule for the instructional program.

As reported in an earlier section, data and impressions gathered at the time of observation also served as a basis for the teacher proficiency classification.

Class scheduling information was gathered to enable the staff to plan observations of all classes.

During the first semester all classes but one were observed by one of the two staff members responsible for the visitations.

During the second semester, 71 of the 74 classes were observed by the three staff members assigned to the task.

Following each observation, the staff member recorded pertinent data and noted any other information which might prove useful.

The observing staff member made suggestions for change to those teachers who were deviating from the program either in content or in their use of instructional materials.

The staff found that many teachers had difficulty in maintaining the pace of one lesson per day. However, a certain pacing latitude was given in the testing whereby the test items incorporated only material covered in the lessons scheduled three days or more prior to the test. Therefore, classes either ahead of schedule or behind schedule as many as 3 lessons were considered on schedule. At the time of the first observations, 46 classes fell within the acceptable range, and all classes fell within the range of 2 lessons ahead of schedule to 15 lessons behind schedule. Eleven teachers were contacted and asked to teach 2 lessons daily to meet the schedule; 5 teachers who were ahead of schedule were asked to make the adjustment necessary to be on schedule.

At the time of the second observation, all teachers with one exception fell within the range of from 3 lessons ahead of schedule to 16 lessons behind schedule. One teacher of an alternate class was 25 lessons behind schedule. In view of the difficulty of remaining on schedule,greater pacing latitude was given in the subsequent testing, so that all classes up to 5 lessons behind schedule would be considered on schedule. 61 classes fell in this range. Where sufficient acceleration was feasible, the project staff contacted the teacher and asked that he teach more than one lesson per day until he met the schedule.

Nine teachers were also interviewed by a member of the project staff. The purpose of the interviews was to determine:

1. The teacher's reaction to the program,
2. The pupils' reaction to the program, and
3. Problem areas in the program.

Although the data gathered was favorable, it could not be related to the objective test findings. A summary of the interview results will be found in Appendix B.

D. The Experimental Model and Its Analysis

1. The model for this experiment as described here differs somewhat from that first proposed. The differences are not substantive. It is rather that the model used in this analysis is a slightly different design although the two designs are of the same class. It was decided that a sex factor should be considered since a large amount of variation in the scores could probably be accounted for by sex differences. This fact was later substantiated by the analysis of the data. The "class within schools" component was eliminated since there was not enough replication of classes within the experimental schools to make this an effective factor. Furthermore, the analysis was simplified by using class or school means rather than individual student scores. This change made the computer analysis simpler and had no other effects of consequence.

2. The final version of the model is a type of hierarchal design. An equal number of schools participating in each experimental method constitute the nesting or hierarchy. The sexes are not nested within the schools because the students of each sex were obtained from each of the schools. Thus the design is partly hierarchal and partly factorial in nature.

3. We apply the symbolism

$$y(m,s,p)$$

to specify the mean score on the y variable made the students of the pth sex in the sth school in which the mth Spanish teaching method was applied. The variation and magnitude of these near scores, in part, can be explained by the factors of the experimental design. That part which cannot be explained is called "error." This is denoted by

$$\acute{\varepsilon}\,(m,s,p).$$

We can write the model in the form

$$\begin{aligned} y(m,s,p) &= \text{constant} + \text{experimental design factors} + \text{error} \\ &= \text{constant} + \text{experimental design factors} + \acute{\varepsilon}(m,s,p) \end{aligned}$$

where the "constant" is a population mean of the yth variable. If the design factors do not explain the magnitude or variation of the mean scores, then

$$y(m,s,p) = \text{constant} + \text{error}.$$

More likely some of the design factors will be helpful. It is the task of the model to determine this.

4. To carry out the analysis, the factors of the design are expressed mathematically as sets of parameters as defined below.

μ = the general mean value of the yth variable obtained by a relevant population of students of which those in this experiment are assumed to be a random sample. This parameter replaces the "constant" term previously used.

$\alpha(m)$ = the contribution to the mean scores of the yth variable produced by exposing the school classes to the mth teaching method. $m = 1, \ldots, M = 9$.

$\alpha(s{:}m)$ = the contribution to the mean score of the yth variable produced by teaching pupils in the sth school by the mth method. $s = 1, \ldots, S = 72$. There are $G = 8$ schools participating in each of the $M = 9$ methods. This parameter set represents hierarchal structure, i.e., schools within methods.

$\alpha(p)$ = the contribution to the mean score of the yth variable produced by teaching pupils of the pth sex. $p = 1, P = 2$.

$\alpha(m,p)$ = the contribution to the mean score of the yth variable produced by teaching pupils with the pth sex by the mth method. This is a simple interaction factor between sex and method.

5. This specification characterizes the design mathematically. We also attempt to explain variation by using auxiliary information called covariance variables. This information consists of ability or pre-test scores. Three of these were used for each analysis made on a y variable. The contribution of these three variables designated as X, W, Z, has the form

$$\beta x(m,s,p) + \gamma W(m,s,p) + \delta Z(m,s,p)$$

where the β, γ, and δ are weights for the variables X, W, and Z and are determined during the analysis of the data. Thus we can write the model for the scores as

$$y(m,s,p) = \mu + \alpha(m) + \alpha(s;m) + \alpha(p) + \alpha(m,p) + \beta X(m,s,p) + \gamma W(m,s,p) + \delta Z(m,s,p) + \varepsilon(m,p,P)$$

Note that error, designated as $\varepsilon(m,s,p)$, represents the effects contributed by the failure of differences between school group mean scores within each method to be the same for each sex.

6. Using this model, a computer program was designed for its analysis and estimation of the data means. A table of the following form was constructed.

Source of Variation		Degrees Of Freedom	Sums Of Squares	Expected Mean Squares
Mean	μ	1	SS(u)	
Methods	$\alpha(m)$	M-1 = 8	SS(M)	$\sigma^2 + P\sigma^2(s:m) + GP\Sigma^2(m)$
Sch/W/Methods	$\alpha(s:m)$	M(G-1) = 63	SS(S:M)	$\sigma^2 + P\sigma^2(s:m)$
Sexes	$\alpha(p)$	P-1 = 1	SS(P)	$\sigma^2 + S\Sigma^2(p)$
Sex by Methods	$\alpha(m,p)$	(P-1)(M-1) = 8	SS(P,M)	$\sigma^2 + G\Sigma^2(m,p)$
Cov X	$\beta X(m,s,p)$	1	SS(X)	
Cov W	$\gamma W(m,s,p)$	1	SS(W)	
Cov Z	$\delta Z(m,s,p)$	1	SS(Z)	
Error	$\varepsilon(m,s,p)$	M(G-1)(P-1)-3 = 60	SS(E)	σ^2
TOTAL	$y(m,s,p)$	GMP = SP = 144	SS(T)	

In the above table, SP = 144 represents the total number of means of the classes which are available in the experiment. The numbers in the column entitled "Degrees of Freedom" indicate the number of parameters which are uniquely estimable in each parameter set. These numbers and the corresponding sums of squares are obtained from the computer solution. The sums of squares indicate the amount of information contributed to the general model by the corresponding parameter set. Each sum of squares is divided by its degrees of freedom to obtain "Mean Squares" (not shown). The expected value of these mean squares will be described subsequently.

7. To proceed further with the objectives of the analysis we must make further assumption about the mean score model other than the structural and mathematical ones so far utilized. We are now obliged to make some statistical assumptions. It is customary, for example, to assume that the c(m,s,p) are normally distributed with mean equal to zero and variance equal to σ^2. The mean square

$$SS(E)/M(G-1)(P-1)-3$$

estimates this variance. The α - type parameter sets may be considered either "fixed" or "random." To explain these terms consider $\alpha(m)$. If the methods are specifically selected or designed for trial in the experiment, we call $\alpha(m)$ a fixed set. If the methods were randomly selected for trial from a population (possibly finite), we call $\alpha(m)$ a random set. If all α - type sets are fixed, we call the model a "fixed model," whereas if all α - types are random, the model is called "random." If some α - types are random, the model is called "random." If some α - types are fixed and others random, then the model is called "mixed."

8. In this experiment $\alpha(m)$ is a fixed set because the methods were specifically designed and not selected at random from a population of methods. Of course, $\alpha(p)$ is fixed since the sexes are determined by nature and we utilize both. Also μ, X, W and Z are considered fixed. Since both $\alpha(m)$ and $\alpha(p)$ are fixed, then $\alpha(m,p)$ is also considered a fixed set.

   However, $\alpha(s:m)$ is considered random, although because of operational circumstances, this is not exactly true. The schools were selected as far as possible by random methods, but inherent in this process is the fact that the teachers within these schools were volunteers for the project and did not represent a random sample. The variance of the $\alpha(s:m)$ is defined as $\alpha^2(s:m)$.

   The term $\Sigma^2(m)$ is the variation of the fixed method effects. This is not a variance in the statistical sense since the methods are not sampled from a population of methods. The other expression, $\Sigma^2(p)$ and $\Sigma^2(m,p)$ are similarly defined. The MEAN and Cov X, W, Z expected mean squares are not specified because they are not required for any important description of this analysis. If the error mean square is compared with a Cov mean square, a significance test of the regression coefficient is obtained.

   The appropriate F tests for significance of the methods, sexes, and sex by method effects are easily obtained by examining the "Expected Mean Squares" column of the analysis of covariance table. The methods are tested for significance by using the mean square for the schools within methods. All other tests use the error mean square. If an effect, say methods, was significant, the adjusted means can be compared pairwise for significance by Duncan's test. It is easy, also, to estimate $\sigma^2(s:m)$ by subtracting the mean square for error from the mean square for schools within methods and dividing this difference by 2.

9. In summary we can say that a mathematical-statistical model was proposed for the Spanish experiment. The components of this model were described in detail, and the analysis procedures were defined. The analysis leads to an analysis of covariance table for a mixed model hierarchal and factorially crossed design. Procedures for making significance tests were described in detail. Estimates of the parameters are output from the computer program used to analyze the data.

E. Test Construction and Testing Procedures

This section of the report contains a description of the instruments which were developed for the measurement and evaluation of pupil achievement in this project. It provides a description of the construction and administration procedures. There were some difficulties in implementing the testing program--see Appendix C.

Objectives

The testing program was constructed to measure pupil achievement in the following areas:

Comprehension of all basic structures

Control of all basic syntactical structures

Ability to produce Spanish sounds.

Achievement was measured through group tests and individual tests.

The group tests consisted of objective multiple-choice items and measured listening comprehension and reading comprehension. These tests were administered to the total pupil population.

The individual tests were comprised of 3 sub-tests. On the 3 sub-tests, the pupil's responses were oral and served as measures of his achievement in:

Vocabulary
Mimicry, and
Oral Reading

Table III illustrates the types of tests administered during each testing period.

Listening Comprehension Tests

Comprehension was defined as the pupil's ability to select the appropriate response from a series of four possible responses in English, when presented with an audible Spanish utterance, within a specified time limit.

The tests were designed to assess vocabulary skills. Instructions and test items for this test were given by tape in an attempt to control the conditions of test administration. The Spanish voices on the test tapes were native speakers of Spanish. The time lapse provided for each pupil response was 20 seconds on the first test (Listening Comprehension Pretest). It was discovered that 20 seconds was more than was generally

TABLE III

TYPES OF TESTS ADMINISTERED AT EACH TESTING PERIOD

TESTS	TESTING PERIOD			
	I (Pre-Test)	II	III	IV
*A. Listening Comprehension	X	X	X	X
*B. Reading Comprehension		X	X	X
**C. Vocabulary		X	X	X
**D. Mimicry (Constructed Response)		X	X	X
**E. Oral Reading		X	X	X

* Group Tests

** Individual Tests - 10% sample (three pupils from each of the 72 classes) needed; therefore, the time lapse was reduced to 10 seconds on subsequent tests. The time lapse allowed for the response did not vary with the length of the stimulus.

1. Stimulus characteristics

   Each test item consisted of a Spanish utterance taken from the Spanish course. The amount of vocabulary tested in each utterance varied because it was not always possible to isolate a single vocabulary item. The length of the utterance was predetermined by the course content. Although utterances were lifted directly from the course content, an attempt was made to limit the length of the utterance to 7 syllables. Each Spanish utterance was spoken twice. These were spoken at what was considered by the staff to be a normal conversational rate of speed.

2. Response characteristics

   The pupil responded by selecting from multiple-choice answers in a test booklet. Pictures and printed alternatives in English were used. However, this combination was used only in the first and second testing periods. No pictures were used in Tests III and IV.

   a. Test items using pictures

      Pictures were used on every item that could be tested by this means. In Test I, such items comprised 50% of the total test. In Test II, picture items comprised 30.2%. There were some significant differences in format between the pre-test (Test I) and Test II.

      (1) Test I (Pre-Test)

      In this test, pupils had 3 choices (1 correct answer and 2 distractors) and were specifically instructed not to make wild guesses, but to leave an answer space blank if they did not know the answer.

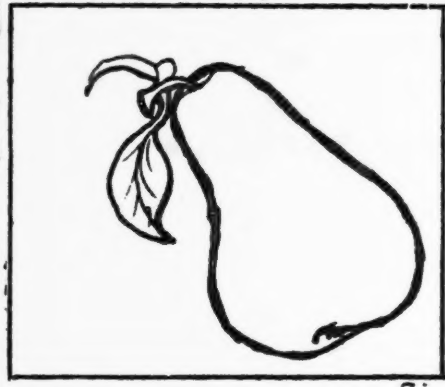
      (2) Test II

      In this test, pupils had 4 choices. The number of choices was increased to reduce the probability that pupils would obtain the correct answer by chance.

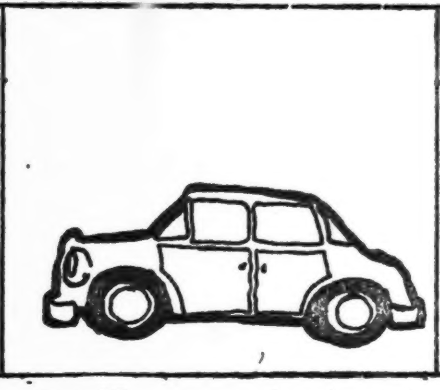
      Sample Item

      An example was first given in English to acquaint the pupil with the format of the testing procedure. The following is a sample item taken from one of the tests. The item assesses the pupil's comprehension of the word "animal". First, the student heard the utterance "ES UN ANIMAL", which was spoken twice. The response choices were:

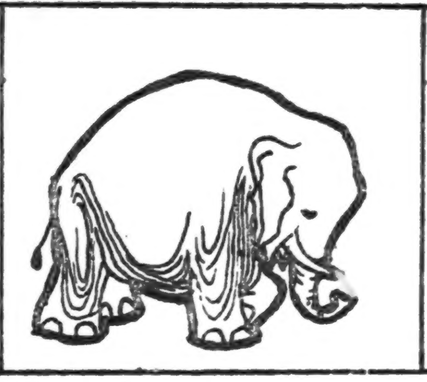
A.



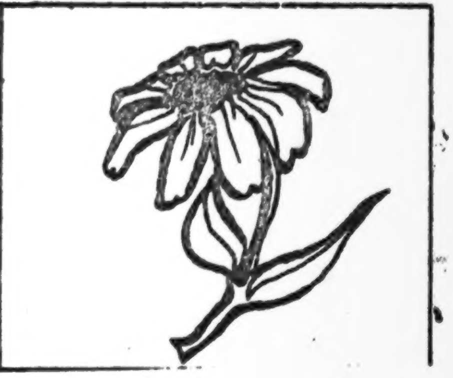
B.



C.



D.



Pupils were instructed to study responses and indicate their choice by marking an answer sheet.* Initially the number of sample items was three. However, this number was reduced to a single sample on subsequent tests as the students became familiar with the format.

b. Test items using printed alternatives in English.

The student heard a Spanish utterance spoken twice. He was then instructed to select the sentence which expressed the same idea. There were four choices, none of which was a "do not know" choice. The choices were constructed in such a way that selection of the correct response was contingent upon the pupil's comprehension of the word or phrase being tested. The other parts of the utterance were held constant whenever possible.

Sample Item

In this example, the pupil's comprehension of the word "ES" was being assessed. The item was "ES PEPE". The pupils heard the utterance in Spanish, spoken twice. Pupils were then asked to study the choices. The four response choices were as follows:

A. I'm Pepe.
B. My name is Pepe.
C. It's Pepe.
D. His name is Pepe.

Pupils were then instructed to indicate their choice by marking the answer sheet.

---

*The Standard Digitek Answer Sheet number DS-1120-A was used on all tests except the pre-test (Test I) and the speaking tests.

Reading Comprehension Tests

The reading comprehension tests followed essentially the same format, and were identical in content to the listening comprehension tests. The test differed in the following respects:

1. The pupil read the Spanish stimulus printed in his test booklet.
2. The time lapse was computed and allowed on the basis of estimated time to read, study the choices, and mark the answer sheet.
3. The items which were used in the listening comprehension were also used in the reading comprehension test. However, the order of these was changed.
4. The same responses developed for the listening comprehension test were used in the responses, but the order in which these choices were listed was changed.

The Speaking Tests

Three speaking tests were developed to assess the pupil's comprehension of Spanish vocabulary and control of Spanish pronunciation. Difficulties in judging the speaking tests are explained in Appendix C .

The speaking tests were individual tests administered to a sample of 3 pupils from each project class.

All instructions to the pupils were pre-recorded on tape to insure uniform conditions in the administration of the tests. All pupil responses were recorded on tape for subsequent evaluation.

The content of each sub-test varied with the test. Descriptions of test content are contained in Table IV.

The test length was limited to 20 minutes per pupil. Because of apprehension and inhibition on the part of the pupil in the testing situation, each project teacher was asked to administer the tests to his own pupils, whenever possible.

1. Speaking Vocabulary Test

This test was designed to see how many of the utterances

TABLE VI

SPEAKING TESTS

TESTS	Description of Content for Each Instructional Period		
	First	Second	Third
Speaking Vocabulary	Covered all basic utterances taught during this period. Thirty-six items.	Covered all basic utterances taught during this period. Thirty-four items.	Covered 2/3 of the basic utterances taught in this period. The other 1/3 of the content consisted of 1/6 from the first I-P,* and 1/6 from the second I-P.* Thirty-six items.
Mimicry (Constructed Response)	Consisted of six couplets based on content from first I-P.*	Consisted of five couplets based on content from second I-P.*	Consisted of five couplets based on content from third I-P.*
Oral Reading	Oral reading of conversation studied in class.	Oral reading of same conversation read for test following first I-P.*	Oral reading of five utterances studied during first and second I-P's,* with some substitutions of lexical items.

*Instructional Period (I-P). This refers to the teaching interval between tests. This interval consisted of 44 twenty-minute teaching periods.

taught in this course could be recalled and pronounced sufficiently well to be comprehensible to a native speaker of Spanish.

(1) Stimulus Characteristics

All stimuli were pre-recorded to insure uniformity of presentation. The pupil heard an expression in English, said twice. The expression was spoken by a native speaker of English at what was considered to be a conversational rate of speed.

The pupil was then allowed time for his response. During the practice sample exercise, the pupil was instructed to express the same idea in Spanish immediately following the English stimulus. No signal was provided to elicit his response.

The object of this test was to ascertain the pupil's ability to recall and articulate the basic utterances taught in the course. The problem then was one of eliciting such responses. Several possibilities were considered and rejected on the grounds that none would insure that the desired response be elicited. Among the ideas rejected were:

(a) The use of pictures

(b) A question and answer situation.

English translations were used because they offered two major advantages:

(a) The students were taught translations

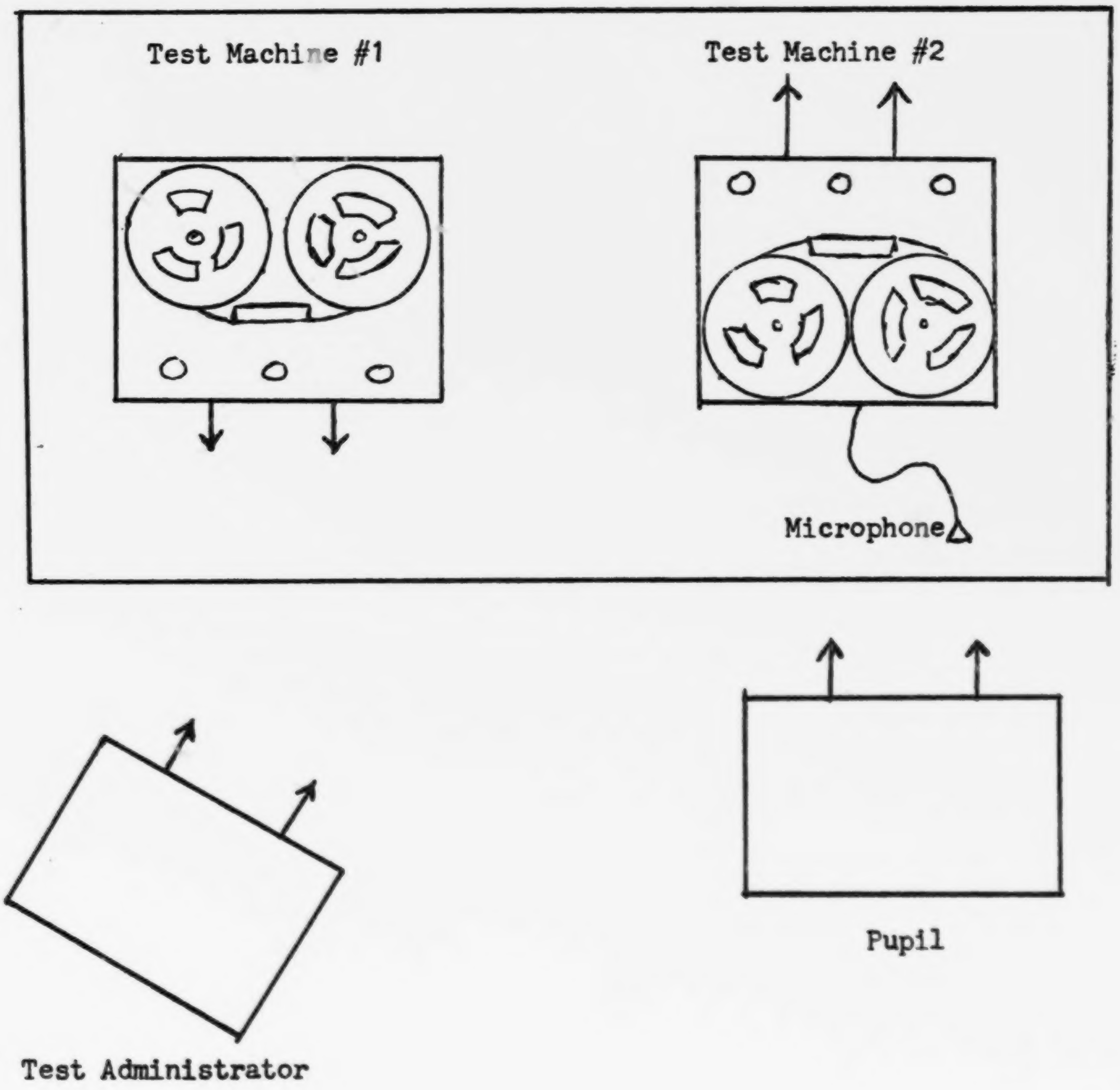
(b) This method would be most likely to elicit the specific desired response.

(2) Response Conditions

Following the taped stimulus, the pupil was provided time to respond. A pause was recorded on the tape to control the time factor and to insure uniformity in administration. After Test I, the length of the pause was standardized at ten seconds. A separate tape recorder was set up to record the pupil responses, as shown in Figure 1.

FIGURE 1

Equipment and Seating Arrangement for Individual Testing


Test Machine #1
Test Machine #2
Microphone
Pupil
Test Administrator

2. Mimicry (Constructed Response) Test

This test was designed to assess the quality of the pupil's pronunciation in a mimicry situation. The test utilized the "question and answer" technique to elicit responses. The questions and the answers were lifted from the content of the course. The questions and answers were spoken by native speakers of Spanish. All pupil responses were recorded on a separate tape recorder.

The format was as follows:

The student heard a question, then he heard the answer. The question was repeated and a chime was sounded to signal the student to repeat the answer.

Example:

English speaker: NUMBER ONE
English speaker: QUESTION
Spanish speaker: ¿CUANTOS AÑOS TIENES?
English speaker: YOU SAY
Spanish speaker: TENGO CATORCE AÑOS.
English speaker: QUESTION
Spanish speaker: ¿CUANTOS AÑOS TIENES?

(CHIME)

The pupil was then expected to respond "TENGO CATORCE AÑOS." A pause was recorded on the taped instructions. The pauses, unlike the pauses in the Vocabulary Test, varied in length. The length of the pause was determined by the length of the utterance.

A major difficulty in testing pronunciation is eliciting the particular utterance containing those elements of phonology to be evaluated. This procedure insured that the pupil respond with the desired utterance which contained the phonological item to be measured. This was not direct mimicry in that the question was repeated and the pupil responded with the answer. The purpose was not to see how well the pupil would imitate the native speaker, but rather to evaluate the quality of his pronunciation of certain sounds at that time.

3. Reading (Oral)

This test was designed to assess the quality of the pupil's pronunciation when the stimulus was a printed word or words. The evaluation was based on the pupil's pronunciation of predefined phonological elements, i.e., consonants and vowels.

The stimulus in all reading tests consisted of material previously studied in the course. It consisted of material which had been practiced orally in class by imitating native speakers of Spanish on tape. There were some major differences in the tests given. The following shows the stimulus material for Tests II-E and III-E:

Pepe y Paco.

Hola.

Hola. ¿Cómo estás?

Bien. ¿Y tú?

*

¡Pepe!

¡Ya voy mamá! ¿Cómo te llamas?

Paco Dorado. ¿Y tú?

Pepe Moreno.

¡Pepe! ¡¡Pepe!! ¡Ven! Date prisa!

Sí, mamá....Mamá me espera. Adiós.

Adiós.

¡Pepe! ¡Date prisa! ¡Corre!

Sí, mamá. ¡Aquí estoy! ¿Qué quieres, mamá?

Stimulus material for TEST III-E:

Pepe y Paco.

Hola.

Hola. ¿Cómo estás?

Bien. ¿Y tú?

*

¡Pepe!

¡Ya voy mamá! ¿Cómo te llamas?

---

*The sentence "Muy bien" was omitted by the typist on both tests.

Paco Dorado. ¿Y tú?

**

¡Pepe! ¡¡Pepe!! ¡Ven! ¡Date prisa!

Sí, mamá....Mamá me espera. Adiós.

Adiós.

¡Pepe! ¡Date prisa! ¡Corre!

Sí, mamá. ¡Aquí estoy! ¿Qué quieres, mamá?

Testing Procedures

The administration of all tests followed procedures prepared by the project staff. These were prepared in booklet form and were explained to the teachers through in-service meetings. Because of the special recording situation required to administer the speaking tests, it was necessary to provide the following:

Diagrams to indicate position of the test administrator, position of pupil, position of tape recorders, as shown in Figure 1.

Detailed instructions on how to thread the tape recorders, where to set the volume, where to set the tone control level, etc.

In-service meetings to demonstrate the above procedures using the actual test materials.

Field Testing

The testing materials used in this study were not field tested individually except for the Pre-Test (Test I). This field test contributed significantly to the development of the final instruments in the following manner:

It established the usability of the format.

It indicated the need for reducing the chance probability in the selection of the correct answer. A step taken to reduce the chance probability in the selection of the correct answer was a directive to the teacher to instruct students not to make wild guesses. It was felt that merely asking pupils not to guess was inadequate. Therefore, the other step taken was to increase the number of choices, thus further reducing the chance probability that the answer was obtained through guessing. No correction for guessing was made in the analysis of the data.

**The sentence "Pepe Moreno" was inadvertently omitted by the typist.

It discriminated between the pupils who spoke Spanish or had previous exposure to Spanish, and the pupils who were non-speakers of Spanish.

It provided data for subsequent rejection of use of pictures for testing purposes beyond Test II. Analysis of the data revealed that no practical difference existed between the use and non-use of pictures. This, coupled with the scarcity of utterances amenable to visualization (35% in Test III), rendered this technique impractical.

Although it is recognized that field testing of all instruments is important, this was not done. The shortage of time and staff personnel precluded field testing the remaining tests.

Scoring Procedures

The group tests were machine-scored, and the pupil's score was the number of correct responses.

The procedures used for scoring the individual tests varied. A description of these procedures follows:

1. Vocabulary test

   For each item the pupil was evaluated on one complete utterance. A pupil was given credit if his response met the following conditions:

   a. The judge recognized what the pupil said.

   b. The utterance was syntactically correct.

   c. The vocabulary was the same as that taught in this course.

2. Pronunciation Tests (Mimicry-Constructed Response)

   Each utterance was studied for the purpose of identifying problems of interference.[1] These problem areas were indicated on a scoring worksheet. Each pupil was judged on his pronunciation of these specific discrete sound elements. The judge heard the pupil's response and marked the scoring worksheet. Responses were judged:

   Right
   Wrong
   A substitution, or
   No response.

---

[1] Interference refers to the problems in pronunciation which occur when a second language is in the process of being acquired.

If a response was judged right, it was not marked. If the response was judged wrong, it was circled. If a response was judged a substitution, the anticipated sound was crossed out and the substitution was written above it in phonetic transcription. Indicated responses omitted by the pupil were considered no response and crossed out.

Sample Item

Utterance: Me gusta el radio.

Me g usta el r a d io.
/g/ /ɹ/ /ae/ /d/

The problems of interference were the following:

/g/ The voiced velar fricative does not exist in English; therefore, there is a strong tendency to substitute the "stop" allophone.[1]

/ɹ/ The voiced multi-vibrant does not exist in English except as a nonsense sound which children sometimes use when mimicking the noise which a machine gune makes when fired.

/ae/ The sound of a as in bat. An intermediate sound between the /ej/ in the English word radio and the /a/ in the Spanish word radio.

/d/ The Spanish pronunciation of the word radio, calls for (rađjo). It was expected that pupils would substitute the voiced, dental stop /d/ for the voiced, dental fricative /đ/. It must be noted that the English orthographic system never used the symbol d to elicit the /đ/ sound. The closest sound to /d/ is the sound of th as in the word "those". It must also be noted that since this word is a cognate, there is a strong tendency to pronounce it the way it was learned in the native tongue.

The number correct for this test was computed by subtracting the number "incorrect", and the number of no responses from the total number of selected phonological items in the test.

The number of items selected for evaluation in each of the Mimicry Constructed Response tests varied. The number of items per test is shown in the tablef following:

---

[1] No attempt was made to drill pupils on these specific sounds. The sounds were accurately and consistently modeled by the native speakers of Spanish on the tapes during the course of the year and were never taught in isolation but in the context of the utterances which comprised the course.

TABLE V

NUMBER OF DISCRETE PHONOLOGICAL ITEMS PER TEST

TEST	II	III	IV
Number of Items	24	17	24

3. Reading Tests

The same scoring procedure which was used to evaluate the pronunciation tests was used to evaluate the reading tests.

In Test II and Test III, each utterance was scored for only one discrete feature of pronunciation. Items were evaluated for correctness. The judge scored an item correct when he heard the expected correct response or a close approximation of it. He scored the item incorrect if he heard the expected error or a close approximation of it. If a pupil failed to respond to an item, it was scored as incorrect. Thus, pupils received credit only for the number of correct responses.

In Reading Test IV, each utterance contained more than one discrete feature of pronunciation. The expected errors, as well as the acceptable correct responses, were identified as in the preceding tests and were scored by the judges in the same way.

F. Chronological Overview

March 1, 1967 - May 31, 1967. The official initiation date of the project was March 1, 1967. Nine days later the Foreign Language Specialist of the San Diego City Schools was designated by the Superintendent as Administrator and Principal Investigator of the project.

Selection of the staff proceeded for approximately one month. Early in April, the two District Resource Teachers for the sixth-grade Spanish program were assigned to the project, and an additional District Resource Teacher was appointed to the project staff. With the appointment of the Coordinator on April 10, 1967, the project became operational.

The major task initially confronting the staff was the analysis of the proposal and the preparation and submission of necessary revisions. From April 10, 1967 to May 1, 1967, daily staff meetings were held for the purpose of clarifying, delimiting, and refining the objectives and the design of the study. In a series of meetings held during April and May, modifications in the experimental design were proposed and analyzed, and detailed plans for implementing the modified design were formulated.

Various methods for the selection of teachers to participate in the project were reviewed by the staff. In recognition of the importance of teacher attitude in such an undertaking, the utilization of teachers who were willing and would volunteer to participate was deemed the most desirable method. The limits of time and project office staffing indicated that this method was the most feasible. Furthermore, in view of the size of the teaching staff, it was felt that there would be no problem in securing enough volunteer teachers for the project. A large number of teachers have been involved in the elementary Spanish program since its initiation. In 1965-66, 152 staff members were listed on the roster of teachers of the elementary foreign language program. Spanish was taught to 225 classes in 100 of the 108 schools offering sixth-grade instruction. In 1966-67, the number of teachers involved was even greater.

Four orientation meetings were conducted by the staff to explain the project to directors, site administrators, and teachers. At these meetings, teachers were invited to participate in the project. Sixty-five teachers from 52 schools volunteered to do so.

The initial steps for the selection of classes were then undertaken. Upon the recommendation of the Director of Testing Services, the following procedure was adopted for this selection:

1. Stratification of the schools of assignment of the teacher volunteers, on the basis of average IQ score obtained on the Lorge-Thorndike test by its sixth-grade pupils in September, 1966, then
2. Assignment to groups through random selection, by number from a random number table.

The staff gathered the data necessary to implement the stratification, although the actual selection was not scheduled until the fall, after teacher assignment changes for the 1967-68 school year would have been effected.

During this period, the staff also initiated the designing of various forms for ultimate use in the analysis and interpretation of data. Drafts of the following forms were developed:

1. A classification of the volunteer group according to proficiency.
2. A check sheet for purposes of describing the activities of a project lesson and the proficiency of the teacher in conducting that lesson.
3. A questionnaire to determine teacher background and attitude.

During the last two weeks of the period, the staff devoted their efforts to various facets of the experimental program. A tentative master calendar and schedule incorporating orientation and in-service sessions were drawn. Work on the analysis and modification of course content and materials was initiated. Pacing charts for the curriculum were developed, and tests to obtain a gross measure of current pupil achievement on the existent Spanish program were written.

June 1, 1967 - August 31, 1967. One of the two major activities requiring special attention of the staff during this period was the continued revision of the experimental design of the project. On the basis of several suggestions for change advanced by staff members and a research consultant of the California State Department of Education, two requests for modification of the approved project were prepared and submitted to the Project Officer of the Office of Education. Upon acceptance of the final request for modification, all schedules and plans for setting the project into operation were revised to reflect these modifications.

The second major undertaking of the period was an analysis of course content and the instructional materials to be used in the project.

The program then in use in the district was a modification of Primer Curso. The publishers' program incorporated reading, whereas the district program was strictly an audiolingual one. A textbook was issued to each teacher for her use, but pupils used pictorial worksheets prepared by the district. These worksheets depicted illustrations from the text and the publishers' visual charts, and contained no printed material in Spanish. In conducting the program, teachers used visual charts, the worksheets, and audiotapes. Also, a set of records was provided each classroom but apparently was not used extensively. The course consisted of 20 minutes of daily instruction.

The evaluation of the Primer Curso content and materials indicated that certain parts of the content presented concepts which were too advanced for most sixth-grade pupils, and the sequence and technique for presenting content on the audiotape was not consonant with the existing San Diego City Schools program. In order to remedy these problems, the staff initiated modification of the course content. This led to a revision and rewriting of the district guide for teaching Spanish in sixth grade, and to the preparation of new tapes to accompany the modified and revised program.

Because the district resource staff was not employed between mid-June and the end of August, the major part of the second undertaking was delayed until after the beginning of the school year. However, in mid-August, one of the district resource staff assigned to the project was employed by the district to assist in revising the guide.

During the first week of June, the staff administered the Primer Curso unit tests to six sixth-grade classes currently studying Spanish. The purpose was to obtain a gross measure of pupil achievement in the regular Spanish program. Scoring of the tests was done, and the analysis was partially completed.

During the summer months, the following clerical operations were performed in preparation for the initiation of the experiment in October:

1. Textbooks for the project classes were purchased, processed and readied for distribution.

2. Tape recorders procured from the district on loan were set aside in sufficient quantity to allow each project teacher a recorder for her own use.

3. Data sheets concerning the site assignments of teachers were reviewed periodically and updated.

4. Data pertinent to the selection of classes such as tables indicating the racial and ethnic distribution by school, and the lists of class types by school were obtained.

September 1, 1967 - November 30, 1967. The major activities of this period, which often took place simultaneously, included:

1. Reprogramming of instructional content
2. Script writing and production of audiotapes
3. Preparation of teachers' guides
4. Planning and creation of evaluative instruments, and
5. Execution of operational procedures, such as the final selection of teachers and classes, the orientation and in-service training of teachers and the supervision of instruction.

Continued operation of the project revealed needs which were not anticipated in the preliminary planning. The need to modify the instructional materials placed unforeseen demands upon the project for clerical help and for personnel to execute the recording of audiotapes. The preparation of the teacher guide for the instructional program was of such magnitude that the district employed two teachers on a part-time basis to supplement the work of the staff. Provision for these demands was not a part of the original proposal, so a reallocation of funds within the budget was effected to make implementation possible.

December 1, 1967 - February 29, 1968. As in the previous three-month period, the staff directed the major part of its effort into revising the teachers' guide, producing audiotapes to accompany the program and writing the tests. Content analyses, lesson plan forms, translations, scripts, and distribution lists were prepared to implement the writing, recording, publishing, and distribution of these materials.

In-service meetings concerning reading instruction and test administration were held for the participating teachers.

Class scheduling data was obtained, and in January and February all project classes were observed. Information concerning the pacing of curriculum in each class was gathered for subsequent use in test score analysis.

Nine teachers, representative of the nine groups in the experiment, were interviewed. The purpose of these interviews was to determine:

1. The teacher's reaction to the program
2. The pupils' reaction to the program, and
3. Problem areas in the program.

In January, at the end of nine weeks of instruction, group and individual tests were administered in all project classes.

March 1, 1968 - May 31, 1968. Additional work was done on the revision of the teachers' guide through March, April, and early May. As in the previous quarters, the magnitude of this task demanded the hiring of additional clerical and professional personnel. In conjunction with this undertaking, comprehensive itemizations of lesson content were prepared. The staff analyzed the previously written guides to identify all entities which could be combined into more meaningful material for review, such as couplets and dialogues.

The writing of tape scripts and preparation of tapes to accompany the program continued throughout the quarter. Various forms and charts to facilitate these operations were developed for office use.

In the latter part of April, a schedule for the second and final visitation of all project classes was prepared. During the month of May, 70 of the 74 participating classes were observed by the three staff members responsible for the task.

Following the visitation series, the staff undertook a classification of all project teachers. The classification was based on:

1. The teacher's mastery of subject
2. Her lesson preparation
3. Her rapport with pupils, and
4. The motivation and participation of pupils.

Two identical in-service meetings concerning reading instruction were held during March. All project teachers who were to initiate reading instruction during the quarter attended one of these meetings.

Additional in-service training was given for 31 project teachers at their request. The staff conducted demonstration lessons and also held conferences to help with any questions or problems concerning the program. In addition, pacing charts and descriptions of the instructional program were prepared and issued to all project teachers. Teachers who had fallen materially behind schedule were contacted by the staff and encouraged to teach extra lessons so that their classes would be on schedule.

As is the district custom, a series of three demonstration lessons was conducted during the school year in each geographical area. Six of the eight demonstration teachers were project teachers. During this period, the final two demonstration lessons of the series were given. The interpretation and discussion following each lesson was, in all cases, conducted by a project staff member.

Early in April, the district invited a group of teachers and principals to evaluate the revised Spanish program. This group, comprised primarily of project teachers, offered suggestions for change and improvement. Where possible, these suggestions were incorporated in the guide revision which was still in process.

Evaluation of the pupil response tapes of the individual test series was begun in mid-March. Three staff members devoted approximately eight days to editing the tapes in preparation for evaluation. The test writer, assisted by other staff members, spent seven days constructing evaluation forms for the tests. During the last week of May, actual evaluation was initiated.

The second and third series of group and individual tests were constructed during the period. The second series was administered to all project classes in mid-March, after 18 weeks of instruction; the third was administered after 27 weeks of instruction, at the end of the year's program.

At teacher request, answer keys for the group tests were provided for all project teachers, so that, if they wished, they might obtain the raw scores of their pupils. Data was compiled, indicating the lesson taught in each class on the day prior to each test series.

At the end of the instructional period, each teacher was asked to submit the following information:

1. The teacher's evaluation of each pupil's progress

2. The names of the pupils who were given Spanish instruction the previous year, and

3. The names of the pupils who are native speakers of Spanish.

Seven schools completed the instructional program and the final testing on May 29, 1968. The remaining 49 schools finished in the succeeding quarter, on June 6, 1968.

June 1, 1968 - August 31, 1968. By June 7, the instructional program and final testing was completed in all project classes. By mid-June, the teacher classification was completed and pupil data submitted by the teachers was compiled. Two district resource teachers were on vacation from mid-June until the end of the project, so the final stage of the project experiment was conducted with a reduced staff. However, to meet deadlines for the completion of test tape editing, the district employed two additional personnel part-time during June and July. During the final quarter, the total attention of the staff was directed toward analysis, evaluation, and interpretation of the data and toward preparing the final report.

CHAPTER III

EXPERIMENTAL RESULTS

Chapter III, following, contains the analysis of the testings. The discussion of the results is in Chapter IV. Tables containing raw-score results for boys and girls and total, for each method, will be found in Appendix D.

A. Listening Test Results

The unadjusted raw score means and the adjusted means for each of the nine methods for the 27-week post testing are shown in Table VI. The analysis of covariance results are presented in Table VII.

Tables VI through XIV concerning the testing results are on the following pages.

TABLE VI.

Raw Score Means and Adjusted Means For Listening Scores

Method		Unadjusted Raw Score Means *	Adjusted Means
One		27.94	25.99
Two		23.63	24.24
Three		25.38	25.25
Four		26.11	26.74
Five		25.17	25.89
Six		26.14	26.47
Seven		25.43	25.44
Eight		26.43	26.09
Nine		25.05	25.15
General Mean		25.6968	
Sex	Male	24.76	25.27
	Female	26.64	26.13

TABLE VII.

ANACOVA Table for Listening Results

Line	Contrast Title	Rank	SSH	MSH	F	Prob	F Denom
1	Mean	1	11.5216	11.5216	0	0	0
2	Cov 1 Y (1.9)	1	37.0774	37.0774	0	0	0
3	Cov 2 Y (3)	1	.5388	.5388	0	0	0
4	Cov 3 X (1) IQ	1	20.9671	20.9671	0	0	0
5	Methods	8	101.5459	12.6932	1.2993	.2603	6
6	Gps /w/ Methods	63	615.4661	9.7693	5.7230	.0000	9
7	Sex	1	21.6835	21.6835	12.7025	.0007	9
8	Sex * Methods	8	14.5630	1.8204	1.0664	.3986	9
9	Sex * Gps /w/ M	60	102.4216	1.7070	0	0	0
10	Total	144	94363.2126	655.3001	0	0	0
11	Total - Mean	143	94351.6910	659.8020	0	0	0.

Post-Test Criterion = 27 wk. Listening Scores.
Pretest Covariate 1 = 9 wk. Listening Scores.
Teacher Rating Cov 2.
IQ Cov 3.

The ANACOVA table for the criterion (Post-Test Listening Scores at 27 weeks) yielded the following items of information:

1. None of the covariates, listening scores at nine weeks, teacher ratings, and IQ's were significant.

---

* Differ slightly from 27-week column totals in Appendix D . Final computer runs made on few less cases but relationships were identical.

2. The differences between methods were not significant.

3. The groups within methods test were highly significant, indicating random sampling of all levels of achievement within a method.

4. The differences between sexes were highly significant.

5. The differences between sexes within methods were not significant.

Because no significant differences between the methods were indicated by the covariance analysis, there was no need to look further at the differences between any two methods.

B. Reading Test Results

The unadjusted raw score means and the adjusted means for each of the nine methods for the 27-week post testing are shown in Table VIII. The analysis of covariance results are presented in Table IX.

TABLE VIII

Raw Score Means and Adjusted Means For Reading Scores

Method		Unadjusted Raw Score Means*	Adjusted Means
One		26.85	22.95
Two		20.49	20.95
Three		25.13	25.33
Four		22.35	24.14
Five		23.48	24.65
Six		26.14	26.84
Seven		23.80	23.56
Eight		24.37	24.26
Nine		23.65	23.58
General Mean		24.0289	
Sex	Male	22.62	23.60
	Female	25.44	24.46

---

* Differ slightly from 27-week column totals in Appendix D . Final computer runs made on few less cases but relationships were identical.

TABLE IX

ANACOVA Table for Reading Results

Line	Contrast Title	Rank	SSH	MSH	F	Prob	F Denom
1	Mean	1	4.8711	4.8711	0	0	0
2	Cov 1 Y (2.9)	1	92.9757	92.9757	4.2683	.0500	0
3	Cov 2 Y (3)	1	2.0497	2.0497	0	0	0
4	Cov 3 X (1) IQ	1	23.2661	23.2661	0	0	0
5	Methods	8	434.0871	54.2609	2.4910	.0204	6
6	Gps /w/ Methods	63	1372.3044	21.7826	9.2467	.0000	9
7	Sex	1	11.2371	11.2371	4.7702	.0329	9
8	Sex * Methods	8	22.1178	2.7647	1.1736	.3300	9
9	Sex * Gps/w/M	60	141.3433	2.3557	0	0	0
10	Total	144	85503.7463	593.7760	0	0	0
11	Total - Mean	143	85498.8752	597.8942	0	0	0

Post-Test Criterion = 27-wk. Reading scores.
Pretest Covariate 1 = 9-wk. Reading scores.
Tchr. Rating Cov. 2.
IQ - Covariate 3.

The ANACOVA table for the criterion (Post-Test Reading Scores at 27 weeks) yielded the following items of information:

1. The pretest reading covariate was significant at the .05 level indicating it was a useful variable with which to control any differences between the groups at the beginning of the experiment.

2. The teacher rating covariate and the ability covariate were not significant.

3. The difference between methods was significant at the .02 level indicating some differences between some of the methods.

4. The groups within methods test was highly significant indicating a random sampling of all levels of achievement within a method.

5. The differences between sexes were significant at the .05 level.

6. The differences between sexes within methods were not significant.

The analysis of variances of this experiment for the reading scores showed that there was a significant difference between the means of the experimental groups. In order to ascertain which groups did differ from each other, a multiple comparison test was made between pairs of means. The basis for this test is given in Reference A† and the procedure is called the Duncan Multiple Range Test. To carry this test out, the means are ordered and differences are computed between every pair. This work is shown below.

TABLE X

Method No.	Mean	Methods Number							
		1	7	9	4	8	5	3	6
		22.95	23.56	23.58	24.14	24.26	24.65	25.33	26.84
2	20.95	2.00	2.61	2.63	3.19	3.31	3.70	4.38*	5.89**
1	22.95		.61	.63	1.19	1.31	1.70	2.38	3.89
7	23.56			.02	.58	.70	1.09	1.77	3.28
9	23.58				.56	.68	1.07	1.75	3.26
4	24.14					.12	.51	1.19	2.70
8	24.26	(Mean differences)					.39	1.07	2.58
5	24.65							.68	2.19
3	25.33								1.51

The Duncan procedure sets up values corresponding to the levels of significance .01 and .05. The mean differences must exceed these values in order for the group pairs to be considered significant. These values are a function of the number of groups and are obtained from tables provided in Reference A. The table values were devised to provide suitable protection against finding wrong significant differences among a set of means.

The application of Duncan's test shows that Group 2 differed at the .01 level of significance from Group 6, and differed at the .05 level of significance from Group 3. These facts are indicated by placing two asterisks (**) after the mean difference 5.89 in the first row of the above table and one asterisk (*) after the mean difference 4.38. It appears, for example, the group which takes oral instruction 20 minutes per day for the first 18 weeks, then oral instruction 15 minutes per day and reading instruction for 5 minutes per day for the remaining 9 weeks (i.e. Group 6) performs in a superior manner on the Reading Test to the group which takes oral instruction for 15 minutes per day for 27 weeks (Group 2). Group 3 performs on the Reading Test in a superior manner to Group 2, but we are slightly less confident about making this assertion. All other pairs of groups are not significant. However, additional trends will be found under discussion of results

---

† Reference A. Duncan, David B., "Multiple Range and Multiple F Tests," Biometrics, Vol. 11, No. 1, 1-42, March 1955.

C. Speaking Test Results

The unadjusted raw-score means and standard deviations for the criterion (27-week Speaking Test No. 3) and the covariate No. 1 (IQ) and covariate No. 2 (9-week Speaking Test No. 3) for each of the nine methods are presented in Table XI.

TABLE XI

Raw Score Data for Speaking Scores
Summary of Raw Score Means and SD's

Method	N	Criterion Mean		Covariate 1*		Covariate 2**	
		Mean	SD	Mean	SD	Mean	SD
1	19	16.00	5.09	4.21	1.32	4.47	2.43
2	13	11.23	5.45	4.61	.83	4.30	2.12
3	18	15.50	5.38	4.33	1.59	4.61	1.89
4	10	9.80	4.60	4.60	1.85	4.10	1.70
5	10	15.00	5.25	4.70	1.61	5.50	1.85
6	18	14.22	5.60	3.94	1.61	4.22	1.51
7	11	11.63	5.51	4.36	2.05	4.63	1.72
8	14	13.78	5.36	4.42	1.44	3.85	2.23
9	17	12.53	3.41	3.47	1.03	3.47	1.72
N =	130	13.59	5.45	4.23	1.53	4.30	2.01

The unadjusted raw score means and the adjusted means for the criterion variable (27-week Speaking Test No. 3) for each of the nine methods are presented in Table XII.

---

* Pupil ability--San Diego Test Performance Scores.

** 9-week speaking test.

TABLE XII

Raw Score Means and Adjusted Means For Speaking Scores

Method	Unadjusted Raw Score Mean	Adjusted Mean
1	16.00	15.76
2	11.23	11.53
3	15.50	15.19
4	9.80	10.35
5	15.00	13.87
6	14.22	14.09
7	11.63	11.32
8	13.78	14.50
9	12.53	12.95
General Mean	13.59	

TABLE XIII

ANACOVA Table for Speaking Test Results Analysis of Covariance Table

Line	Contrast Title	Rank	SSH	MSH	F	Prob
1	Mean	1	---	---	---	---
2	Cov 1 (Ability)	1	---	---	---	---
3	Cov 2 (9-wk.pretest)	1	---	---	---	---
4						
5	Methods	8	477.72	59.71	2.21	.05
6	Grps/w/Methods	119	3238.69	27.21	---	---
	Total	130	3716.41			

The analysis of covariance indicated there was a significant difference between methods at the .05 level. In order to ascertain which groups differed from each other a T-test was computed for each pair of means. Table XIV contains the results of this analysis.

The adjusted means in Table XIV are ordered from lowest to highest mean (by method). The intersections of the matrix indicate the actual differences between any two means.

TABLE XIV

Method No.	Mean	4	7	2	9	5	6	8	3	1
		10.35	11.32	11.53	12.95	13.87	14.09	14.50	15.19	15.76
4	10.35		.97	1.18	2.60	3.52	3.74	4.15	4.84*	5.41**
7	11.32			.21	1.63	2.55	2.77	3.18	3.87	4.44*
2	11.53				1.42	2.34	2.56	2.97	3.66	4.23*
9	12.95					.92	1.14	1.55	2.24	2.81
5	13.87						.22	.63	1.32	1.89
6	14.09							.41	1.10	1.67
8	14.50	(mean differences)							.69	1.26
3	15.19									.57
1	15.76									

* .05.
** .01.

The application of the T-tests indicates that Method One differed from Methods Two and Seven at the .05 level and from Method Four at the .01 level. Method Three differed from Method Four at the .05 level. No other significant differences were uncovered.

CHAPTER IV

DISCUSSION OF THE RESULTS

A. Hypotheses

Concerning no methods differences in listening, reading, and speaking test achievement, the null hypothesis was accepted for listening, and rejected at the .02 level of significance for reading, and the .05 level of significance for speaking, indicating there were significant differences in methods for reading and speaking.

The hypothesis concerning no differences in teacher effect on pupils using materials prepared for the project was accepted--the teacher rating covariate was not statistically significant in listening and reading, indicating that success was obtained in making the effect on teaching materials independent of teacher differences. The alternative would be that teachers presented the material as planned for the project. It would be indicative but not conclusive that the materials were independent of teacher ability.

The hypothesis concerning no sex differences in listening and reading test achievement was rejected for total achievement over all methods, only. Over all methods, girls achieved better than boys, with statistical significance being at the .01 level in listening, and the .03 level in reading. No statistically significant differences between boys and girls existed in listening and reading between methods, or groups within methods.

B. Methods Differences

While there were statistically significant differences in methods for both reading and speaking, these differences are of interest in terms of only one or two of the nine methods in each area, and because of conflicting data they would not seem to be of practical value.

In reading, for instance, while Methods 3 and 6 were superior to Method 2, Methods 3 and 6 were not superior to Methods 1, 4, 5, 7, 8, and 9; nor did these latter six methods differ statistically from Method 2. The same situation existed in the area of speaking.

However, if trends were to be considered insofar as any score differences are concerned (without consideration of statistical significance), there was revealed a rather consistent pattern of methods that differed in achievement from the remaining methods. A ranking, in both reading and speaking, by the number of methods exceeded by each method in raw score is shown in Table XV. Also given in Table XV is the method description. From this table, the following observations can be made:

1. In reading, trends revealed that prereading oral or audio-lingual instruction followed by a combination of oral and reading instruction produced better results than the method

TABLE XV

Comparisons on Raw Score Differences By Number of Methods Exceeded by Each Method

Reading			Speaking		
Meth.	No. Methods With Mean Score Lower	Description	Meth.	No. Methods With Mean Score Lower	Description
6	8	18W Oral 20' 9W Oral 15'/Rdg 5' (Curr)	1	8	27W Oral 15'/Rdg 5' (Curr)
3	7	9W Oral 20' 18W Oral 15'/Rdg 5' (Curr)	3	7	9W Oral 20' 18W Oral 15'/Rdg 5' (Curr)
5	6	9W Oral 20' 18W Oral 15'/Rdg 5' (Rev.)	8	6	18W Oral 20' 9W Oral 15'/Rdg 5' (Rev.)
8	5	18W Oral 20' 9W Oral 15'/Rdg 5' (Rev.)	6	5	18W Oral 20' 9W Oral 15'/Rdg 5' (Curr)
4	4	9W Oral 20' 18W Oral 15'/no reading	5	4	9W Oral 20' 18W Oral 15'/Rdg 5' (Rev.)
9	3	27W Oral 20'/no reading	9	3	27W Oral 20'/no reading
7	2	18W Oral 20' 9W Oral 15'/no reading	2	2	27W Oral 15'/no reading
1	1	27W Oral 15'/Rdg 5'	7	1	18W Oral 20' 9W Oral 15'/no reading
2	0	27W Oral 15'/no reading	4	0	9W Oral 20' 18W Oral 15'/no reading

beginning with a combination of oral and reading and the methods where no reading instruction was given at all. However, the number of weeks of prereading oral instruction that was most productive was not determinable--either nine or eighteen weeks was just as effective.

Trends would appear to indicate that regardless of the number of weeks of prereading oral instruction, current rather than review material should be used.

While quite tenuous, in reading 18 weeks of prereading oral instruction would seem to be indicated when using current content; and nine weeks of prereading oral instruction when using review materials. In speaking, no pattern emerged, except that of the highest of four ranked methods, current material was associated with three.

2. In both reading and speaking, Methods 2, 4, 7, and 9 ranked lowest in criterion test achievement; and all consisted of oral work only.

3. In reading, Methods 3 and 6 ranked highest in exceeding other methods by one or more raw score units and both used current reading material in their methods. Method 6 was clearly superior in exceeding all other methods by two or more raw-score units in criterion test achievement.

4. In speaking, Methods 1 and 3 ranked highest in exceeding other methods by one or more raw-score units. Both used current material. Methods 6 and 8 were equal in the number of methods exceeded by one or more raw-score units--since one was using current reading material and the other review material, no difference in materials was of any practical significance.

CHAPTER V

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

The main objective of this study was to determine whether the inclusion of reading instruction increased achievement in reading, speaking, and listening in a program of beginning Spanish at the sixth-grade level; and if so, how much. To answer this question, nine methods were designed to present pupils with varying amounts of reading instruction in a 27-week period. These methods contained combinations of oral instruction only preceding oral instruction plus reading.

A. Conclusions

From the standpoint of statistical significance only, the results of this study were inconclusive.

. . . The various methods produced no significant differences in listening achievement.

. . . Although statistically significant differences in reading and speaking occurred for different methods, these differences were in conflict; that is, two similar methods which were each superior to other methods were not superior to the same method.

If, however, the data is examined for indications other than those which comply with strict statistical significances, then certain interesting trends are apparent.

. . . Reading and speaking achievement were higher in those methods in which reading instruction was part of the method. The presence of reading instruction in the upper 50 per cent of methods but not in the lower 50 per cent indicates that reading may be a discriminating factor.

. . . Reading achievement was higher when oral instruction only preceded a combination of oral and reading instruction, but the data gives no indication of the optimum amount of this oral instruction.

. . . Speaking achievement seemed to be independent of whether the method began with a combination of oral and reading instruction or with a period of oral instruction preceding a combination of oral and reading instruction.

B. Implications

A generally accepted tenet of the audiolingual approach to foreign language learning is that new material should generally be introduced orally before it is practiced through reading, especially at the beginning levels of instruction. Although the results of this study do not provide statistically significant support to this tenet, neither do the results indicate that the tenet should be rejected. In fact, when the data are evaluated from a non-statistically significant point of view, there are indications that

1. The inclusion of reading in a program of Spanish instruction for sixth-graders results in not only higher achievement in reading but in speaking as well.

2. The provision of a prereading period, i.e., oral instruction alone preceding a combination of oral and reading instruction--results in a higher achievement in reading.

Since achievement in speaking seemed in this study not to be dependent on the inclusion or non-inclusion of a prereading period, the general implication is that the curriculum probably should continue to provide for a period of purely oral foreign language instruction at the sixth-grade level. Following this prereading period (and for a duration not indicated by the study) instruction should probably include reading in combination with oral instruction.

C. Recommendations

It is obvious that the findings of this study apply only to a particular body of Spanish materials used with a particular group of children under a particular set of circumstances. To generalize beyond the conclusions arising out of this project, even if the results showed statistical significance, many more studies would have to be conducted using different materials with different populations under different conditions. Thus, one major recommendation arising from the study is that this be done.

In addition, it is suggested that future studies consider the following recommendations:

1. The amount of reading which is included in the combination of oral and reading instruction should be more than five minutes in a total of 20 minutes of instruction.

   In this study, the length of time devoted to reading in relation to the total time of instruction was so short that the nature of the prereading instruction was not significantly different from the nature of the instruction which included reading.

2. Each of the two major questions investigated in this study should in any future investigations be incorporated into a separate research design. That is, the question of whether the inclusion of reading in a foreign language program affects achievement should be investigated in a design separate from an investigation of how much reading instruction results in maximum achievement.

   In this study, the amount of reading instruction was reciprocal to the amount of prereading instruction. That is, in the 27 weeks during which the study was conducted, if the prereading period (oral instruction only) was 27 weeks, there was no reading; if the prereading period was 18 weeks, there were 9 weeks of reading; 9 weeks of prereading, 18 weeks of reading; no prereading, 27 weeks of reading. But in relation to the total daily allotment of time, the reading component was always a fixed amount of time. The second question might better have been answered if there had been a gradient in the amount of time allotted daily to different methods tested.

3. In the selection of materials for studies involving the relationship of reading instruction to achievement in reading and other skills, the content should present a sequence which progresses in difficulty from simple to complex and in which later content is dependent upon preceding content.

   Because the materials used in this study did not present such a sequence—i.e., what was studied in the fifth week was not necessarily dependent upon what had preceded, and the content of the sixteenth week might readily be less complex than that of the first week—there was little need to conduct the study for the entire 27-week period. Inspection of data gathered at the 9- and 18-week benchmarks indicates a confirmation of the results at the end of the study.

In general, the major recommendation growing out of this study is that there is great need for further investigation of the influence of reading instruction upon achievement in reading, speaking, and listening in beginning foreign language learning.

CHAPTER VI

SUMMARY

In summary, the project was designed to answer two primary questions:

1. Does the inclusion of reading in a program of sixth-grade beginning Spanish instruction affect achievement in the skills of listening, speaking, and reading?
2. If the inclusion of reading results in a significant difference in achievement, what amount of reading instruction results in the most significant difference?

The study was also designed to answer a secondary question:

> Does a significant difference in achievement take place when reading instruction utilizes a combination of new and review content rather than content which is entirely review?

Nine groups, each composed of eight classes, participated in the study. Each group received daily instruction in Spanish, varying from 15 to 20 minutes. The district modification of Primer Curso served as the course of study for all groups.

The amount of content was held constant in all nine methods, but the lengths of time devoted to reading instruction and the lengths of time devoted to oral instruction were varied. Content used for reading instruction was varied also.

The duration of the dat[illegible]ion period was 27 instructional weeks.

The teachers who participated were selected from the group of teachers teaching Spanish at that time, who volunteered to participate.

In order to obtain a representative sample of the pupils, a stratified random sampling procedure was used. Each method by this procedure contained approximately the same range of ability among its groups as any other method.

The dependent variables used as criteria were listening, reading and speaking tests administered at the close of the 27-week period. Analysis of covariance was used for the analysis. Independent covariates used as controls for each criterion variable were listening, reading, and speaking pretests, respectively, at the 9-week period; teacher ability ratings, and scholastic ability of pupils were controls in each comparison. The experimental design was partly hierarchial and partly factorial in nature involving sex, class groups and methods.

Concerning no methods differences in listening, reading, and speaking test achievement, the null hypothesis was accepted for listening, and rejected at the .02 level of significance for reading, and the .05 level of significance for speaking, indicating there were significant differences in methods for reading and speaking.

From the standpoint of statistical significance only, the results of this study were inconclusive.

. . . The various methods produced no significant differences in listening achievement.

. . . Although statistically significant differences in reading and speaking occurred for different methods, these differences were in conflict; that is, two similar methods which were each superior to other methods were not superior to the same method.

If, however, the data is examined for indications other than those which comply with strict statistical significance, then certain interesting trends are apparent.

. . . Reading and speaking achievement were higher in those methods in which reading instruction was part of the method. The presence of reading instruction in the upper 50 per cent of methods but not in the lower 50 per cent indicates that reading may be a discriminating factor.

. . . Reading achievement was higher when oral instruction only preceded a combination of oral and reading instruction, but the data gives no indication of the optimum amount of this oral instruction.

. . . Speaking achievement seemed to be independent of whether the method began with a combination of oral and reading instruction or with a period of oral instruction preceding a combination of oral and reading instruction.

The conclusions and implications of this study should be confined to the specific instructional program in use in this project. They should not be generalized or applied to other foreign language courses of study.

For more conclusive results in this area, it is recommended that further studies be made.

APPENDIX A

IQ Ranges and Means for Boys and Girls And Total for Each Method (1-9)

COOPERATIVE SPANISH STUDY

ABILITY

		Method 1	Method 2	Method 3	Method 4
Range					
Boys	SDTPS	2.3-5.2	3.0-6.2	2.3-6.2	3.6-6.7
	IQ	122-99	117-91	122-91	112-87
Girls	SDTPS	2.9-512	2.6-6.6	2.2-6.9	4.2-6.9
	IQ	117-99	120-88	123-85	107-85
Total Group Means					
Boys	SDTPS	4.46	4.25	4.80	5.38
	IQ	104.8	106.5	102.1	97.5
Girls	SDTPS	4.29	4.55	4.67	4.91
	IQ	106.2	104.1	103.1	101.2
Total	SDTPS	4.38	4.41	4.74	5.16
	IQ	105.5	105.2	102.6	99.2

APPENDIX A

IQ Ranges and Means for Boys and Girls And Total for Each Method (1-9)

COOPERATIVE SPANISH STUDY

ABILITY

		Method 5	Method 6	Method 7	Method 8	Method 9
Range						
Boys	SDTPS	4.1-7.0	4.5-6.1	3.4-6.9	3.1-6.6	3.0-6.3
	IQ	108-84	104-92	113-85	116-88	116-90
Girls	SDTPS	3.5-6.6	4.2-5.8	3.3-6.8	3.4-6.7	2.5-6.3
	IQ	112-88	107-94	114-86	113-87	120-99
Total Group Means						
Boys	SDTPS	5.29	5.02	4.68	4.80	4.55
	IQ	98.2	100.4	103.1	102.1	104.1
Girls	SDTPS	5.26	4.77	4.68	4.65	4.16
	IQ	98.4	102.4	103.1	103.3	107.2
Total	SDTPS	5.27	4.88	4.68	4.73	4.37
	IQ	98.3	101.5	103.1	102.6	105.6

APPENDIX B

Interviews with Teachers

At the end of the first semester, a staff member interviewed nine project teachers. All nine groups were represented in this selection, and all IQ strata (by schools, as indicated by sixth-grade pupil achievement on the Lorge-Thorndike test, September, 1966), with the exception of the IQ 101-102 stratum, were represented.

The purpose of these interviews was to determine:

1. The teacher's reaction to the program

2. The pupils' reaction to the program, and

3. Any problem areas in the program.

The data gathered by the interviewer indicated that:

a. The lesson plans were excellent, well planned, and functional, though one teacher found the plans confining. Suggestions were made to incorporate more use of visuals, more new material, more vocabulary items in categories such as numbers, colors, and clothing, and more teaching of materials in couplets.

b. The tapes were useful and superior to those used in previous years. One teacher felt that the short utterances were repeated too many times. Three teachers found the pauses either too long or too short. One teacher felt the new material was, at times, too difficult to understand.

c. The pupils' reaction to the tapes and/or the program was favorable, according to 8 teachers. One teacher claimed that his pupils did not respond negatively. One teacher said that her pupils were bored with the short phrases. The suggestion to use the backward build-up on the tape was made.

d. Seven of the 9 teachers regarded themselves as adequately prepared or more than adequately prepared to conduct the program. One teacher claimed she had no self-confidence although she felt her background was adequate. One teacher felt that his self-confidence was increasing.

e. All 9 teachers felt that Spanish should continue to be included in the elementary curriculum.

During the second semester, the district invited a group of teachers and principals to evaluate the Spanish program. In their meeting, 5 project teachers, as members of the committee, presented their opinions and suggestions.

In general, they felt the program was worthwhile, though the necessity to adhere strictly to the prescribed plan was often burdensome. The following suggestions for improvement were made:

1. Add more variety and more time to the warm-up segment of the lesson,
2. Build more visual use and interrelatedness of content into the review segment,
3. Equalize the content load in the lessons,
4. Add more variety in the activities, and
5. Increase amount of content for more able pupil groups.

Two other concerns were mentioned. The manipulation of several visuals at once is difficult for some teachers; for these teachers perhaps charts would be more convenient. Secondly, the project reading programs required too much writing on the part of the teacher. In that the text has little material for the pupils to read, sentence, strips, work sheets or similar types of materials must be prepared by the teacher for pupil use.

APPENDIX C

Problems in Implementing the Testing Program

There were some problems in implementing the testing program. The individual testing program for each instructional period required at least 60 minutes of teacher time. This presented difficulties in schools in which personnel were not available to supervise the classes while the individual tests were administered by the teachers.

In addition, two testing calendars were necessary to offset scheduling differences which occurred when participating classes suspended the Spanish program for the week which they spent at sixth-grade camp. The following schedule for Test III illustrates the type of calendar used.

A. All classes which have not attended camp during the October 16 - March 15 period take tests in the following order and on the dates listed below:

	Group Tests	Dates
1.	Listening Comprehension Test, III-A	Monday, March 18, 1968
2.	Reading Comprehension Test, III-B	Tuesday, March 19, 1968

	Individual Tests	Dates
1.	Speaking Vocabulary Test, III-C	March 18 or 19, 1968
2.	Pronunciation Test, III-D	March 18 or 19, 1968
3.	Oral Reading Test, III-E	March 18 or 19, 1968

B. All classes which attended camp during the October 16 - March 15 period take tests in the following order and on the dates listed below:

	Group Tests	Dates
1.	Listening Comprehension Test, III-A	Monday, March 25, 1968
2.	Reading Comprehension Test, III-B	Tuesday, March 26, 1968

	Individual Tests	Dates
1.	Speaking Vocabulary Test, III-C	March 25 or 26, 1968
2.	Pronunciation Test, III-D	March 25 or 26, 1968
3.	Oral Reading Test, III-E	March 25 or 26, 1968

APPENDIX C

Difficulties in Judging Speaking Tests

An inherent difficulty in judging speaking tests was the problem of judging them objectively. This concern resulted in the development of several measures designed to achieve greater objectivity in the scoring of these tests. They are as follows:

(a) The judges were involved in the development of the scoring sheet.

(b) The conditions under which responses would be acceptable were defined regarding grammar, syntax, pronunciation, and completeness. Several trial examples were developed for the purpose of clarification.

(c) The response was to be expressed utilizing the vocabulary as taught in the course. For example, if the expected response was "Lavo los platos," and the pupil responded with "Lavo los trastes," he was not given credit. The word "trastes," although a suitable substitution, was not taught in this course.

(d) All responses were recorded on tape. This enabled the judge to review the response in cases where there was doubt.

(e) The scoring sheet was designed to aid the memory of the judge. Each scoring sheet had the expected response printed on it. As the pupil responded, the judge circled omissions, inserted additions; and in cases of substitutions, the word from the expected response was crossed out; and the substitution was written above it. If the student did not respond, the judge simply checked a "No Response" column.

In the following examples, some acceptable responses will be given to illustrate two acceptable types of responses.

Sample Item

Stimulus: "My name is Pepe."

Expected response: "Me llamo Pepe."

Acceptable response: "Me llama Pepe." (Grammatical error)
"Me llamo es Pepe." (Syntactical error)

Incomplete responses were unacceptable if there was an omission of the article, or if meaning was lost.

Sample Item

Stimulus: "How old is Belita?"

Expected response: "¿Cuántos años tiene Belita?"

Unacceptable responses: "¿Cuántos....Belita?"

"¿....años Belita?"

Acceptable response: "¿Cuántos años....Belita?"

However, note the difference in the following sample item.

Stimulus: "I wash the dishes."

Expected response: "Lavo los platos."

Unacceptable responses: "...los platos."

"...platos."

"Lavo..."

The above expressions were devoid of meaning, hence, unacceptable.

Acceptable responses: "Lavo los platos."

"Lavo...platos."

APPENDIX D

Summary of Mean Raw Scores
For Each Method

COOPERATIVE SPANISH STUDY*

Listening (N = 1,901)

	Y (1,0) 0 Weeks	Y (1,9) 9 Weeks	Y (1,18) 18 Weeks	Y (1,27) 27 Weeks
Total Method 1 (N = 237)				
Boys	7.63	37.85	26.24	26.33
Girls	7.46	38.97	28.46	28.37
Total Boys and Girls	7.54	38.35	27.28	27.29
Total Method 2 (N = 230)				
Boys	7.16	34.75	23.54	23.08
Girls	7.67	36.63	24.47	24.41
Total Boys and Girls	7.41	35.70	24.00	23.73
Total Method 3 (N = 205)				
Boys	8.59	36.71	25.30	24.77
Girls	8.54	37.00	27.84	25.62
Total Boys and Girls	8.57	36.85	26.56	25.19
Total Method 4 (N = 200)				
Boys	6.93	34.90	21.94	24.74
Girls	7.05	37.59	23.65	27.20
Total Boys and Girls	6.98	36.13	23.63	25.85
Total Method 5 (N = 186)				
Boys	8.86	34.75	23.39	23.73
Girls	9.71	35.68	25.38	25.35
Total Boys and Girls	9.35	35.25	24.46	24.61
Total Method 6 (N = 213)				
Boys	7.60	33.61	22.39	23.95
Girls	8.39	38.22	27.20	27.33
Total Boys and Girls	8.05	36.08	24.92	25.76
Total Method 7 (N = 222)				
Boys	8.80	34.32	22.69	23.57
Girls	8.72	37.54	24.83	26.22
Total Boys and Girls	8.76	35.77	23.65	24.88
Total Method 8 (N = 196)				
Boys	9.31	37.31	24.87	25.68
Girls	8.88	38.79	26.49	27.12
Total Boys and Girls	9.10	38.03	25.66	26.38
Total Method 9 (N = 213)				
Boys	6.54	36.09	24.24	24.36
Girls	7.48	37.07	26.04	25.15
Total Boys and Girls	6.96	36.54	25.07	24.72

* Teacher Rating and Pupil Ability are the same as listed under Reading for each method.

APPENDIX D

Summary of Mean Raw Scores For Each Method

COOPERATIVE SPANISH STUDY

Reading (N = 1,893)

	Y (2,9) 9 Weeks	Y (2,18) 18 Weeks	Y (2,27) 27 Weeks	Y (3) Tchr.Rtg.	X (1) Ability*
Total Method 1 (N = 233)					
Boys	35.69	52.65	24.58	1.52	4.46
Girls	37.17	56.42	27.04	1.47	4.29
Total Boys and Girls	36.41	54.44	25.73	1.50	4.38
Total Method 2 (N = 229)					
Boys	31.41	46.69	19.95	1.96	4.25
Girls	34.03	49.40	21.68	1.90	4.55
Total Boys and Girls	32.70	48.03	20.79	1.93	4.41
Total Method 3 (N = 206)					
Boys	32.69	49.92	23.79	1.79	4.80
Girls	34.11	53.27	26.00	1.89	4.67
Total Boys and Girls	33.39	51.60	24.90	1.84	4.74
Total Method 4 (N = 195)					
Boys	30.28	47.09	20.17	2.42	5.38
Girls	33.72	52.46	24.42	2.50	4.91
Total Boys and Girls	31.83	49.48	22.09	2.46	5.16
Total Method 5 (N = 185)					
Boys	30.64	47.02	20.75	1.69	5.29
Girls	32.57	50.36	23.43	1.71	5.26
Total Boys and Girls	31.69	48.82	22.20	1.70	5.27
Total Method 6 (N = 217)					
Boys	29.59	44.65	22.71	2.16	5.02
Girls	35.46	53.76	27.06	2.05	4.77
Total Boys and Girls	32.71	49.52	25.07	2.10	4.88
Total Method 7 (N = 215)					
Boys	31.81	46.08	21.97	1.51	4.68
Girls	34.75	50.60	24.58	1.54	4.68
Total Boys and Girls	33.15	48.12	23.24	1.52	4.68
Total Method 8 (N = 197)					
Boys	33.03	50.88	23.25	1.95	4.80
Girls	35.32	52.83	25.29	1.99	4.65
Total Boys and Girls	34.14	51.83	24.25	1.97	4.73
Total Method 9 (N = 216)					
Boys	32.17	48.22	21.62	2.06	4.55
Girls	33.66	52.13	23.92	2.06	4.16
Total Boys and Girls	32.85	50.02	22.66	2.06	4.37

* San Diego Test Performance Scores.

APPENDIX D

Summary of Means for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 1 Listening

Group	Sex	Y (1,0) 0 Weeks	Y (1,9) 9 Weeks	Y (1,18) 18 Weeks	Y (1,27) 27 Weeks
1	Boys	7.46	30.05	28.11	23.18
	Girls	8.92	27.30	28.95	24.86
2	Boys	8.33	41.00	30.00	31.00
	Girls	7.53	40.12	31.62	31.50
3	Boys	6.25	34.63	24.54	20.91
	Girls	7.41	35.82	26.50	21.89
4	Boys	5.53	41.55	31.89	31.05
	Girls	7.14	41.77	32.77	33.35
5	Boys	7.53	35.64	20.93	22.00
	Girls	6.46	37.00	26.07	26.06
6	Boys	8.25	37.77	26.07	27.15
	Girls	5.55	40.56	28.60	29.47
7	Boys	6.50	42.50	31.18	32.30
	Girls	6.14	44.14	33.00	31.72
8	Boys	11.57	38.81	21.87	23.22
	Girls	12.00	41.90	28.70	29.84
Total	Boys	7.63	37.85	26.24	26.33
	Girls	7.46	38.97	28.46	28.37
Total Boys and Girls		7.54	38.35	27.28	27.29

APPENDIX D

Summary of Means for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 1
(cont'd)

Reading

Group	Sex	Y (2,9) 9 Weeks	Y (2,18) 18 Weeks	Y (2,27) 27 Weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys	27.27	44.07	27.84	2.00	4.47
	Girls	30.37	44.81	30.21	2.00	4.33
2	Boys	38.18	58.50	28.31	1.00	4.20
	Girls	38.93	62.75	30.55	1.00	4.35
3	Boys	32.00	49.36	18.69	1.00	4.33
	Girls	34.44	52.77	21.66	1.00	3.89
4	Boys	41.41	63.05	31.85	1.00	2.30
	Girls	41.08	64.46	33.88	1.00	2.88
5	Boys	32.57	42.20	20.29	1.00	4.41
	Girls	35.91	51.92	26.87	1.00	4.37
6	Boys	36.69	52.15	26.53	2.00	5.00
	Girls	40.06	55.43	27.23	2.00	4.25
7	Boys	41.00	62.37	30.95	1.00	4.44
	Girls	40.28	66.14	30.45	1.00	3.80
8	Boys	36.25	46.12	20.77	2.00	5.25
	Girls	39.80	57.00	26.15	2.00	5.16
Total	Boys	35.69	52.65	24.58	1.52	4.46
	Girls	37.17	56.42	27.04	1.47	4.29
Total Boys and Girls		36.41	54.44	25.73	1.50	4.38

*** San Diego Test Performance Scores.**

APPENDIX D

Summary of Means for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 2 **Listening**

Group	Sex	Y (1,0) 0 Weeks	Y (1,9) 9 Weeks	Y (1,18) 18 Weeks	Y (1,27) 27 Weeks
1	Boys	8.95	36.36	28.65	26.60
	Girls	8.14	39.66	30.40	29.19
2	Boys	8.09	37.85	24.26	21.23
	Girls	8.93	36.56	26.06	27.41
3	Boys	5.26	35.25	25.87	22.58
	Girls	5.93	41.00	26.87	24.79
4	Boys	4.00	30.93	14.75	19.82
	Girls	4.75	34.31	21.40	22.43
5	Boys	7.18	35.54	22.00	22.73
	Girls	4.33	36.93	23.38	24.35
6	Boys	7.11	32.00	23.11	22.54
	Girls	7.82	34.50	21.77	21.23
7	Boys	9.78	34.84	20.93	23.55
	Girls	8.77	32.71	19.85	21.58
8	Boys	6.36	31.10	23.18	23.58
	Girls	15.37	38.43	27.43	26.12
Total	Boys	7.16	34.75	23.54	23.08
	Girls	7.67	36.63	24.47	24.41
Total Boys and Girls		7.41	35.70	24.00	23.73

APPENDIX D

Summary of Means for Each Group
Within a Method

COOPERATIVE SPANISH STUDY

Method 2
(cont'd)

Reading

Group	Sex	Y (2,9) 9 Weeks	Y (2,18) 18 Weeks	Y (2,27) 27 Weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys	36.26	53.87	24.39	2.00	2.95
	Girls	38.28	57.00	27.94	2.00	2.56
2	Boys	31.84	48.73	20.66	2.00	4.20
	Girls	35.06	52.12	26.64	2.00	4.50
3	Boys	31.25	47.93	18.47	1.00	3.94
	Girls	37.56	53.75	22.89	1.00	4.50
4	Boys	27.85	33.93	15.17	3.00	5.14
	Girls	32.20	42.56	16.56	3.00	4.60
5	Boys	32.92	44.46	22.40	1.00	4.69
	Girls	34.57	47.64	23.75	1.00	4.46
6	Boys	27.55	47.89	18.36	2.00	5.40
	Girls	30.77	45.05	17.62	2.00	5.21
7	Boys	28.53	44.07	18.65	2.00	6.21
	Girls	28.57	43.71	18.26	2.00	4.93
8	Boys	29.90	47.54	17.83	3.00	5.66
	Girls	37.83	57.57	19.57	3.00	6.62
Total	Boys	31.41	46.69	19.95	1.96	4.25
	Girls	34.03	49.40	21.68	1.90	4.55
Total Boys and Girls		32.70	48.03	20.79	1.93	4.41

* San Diego Test Performance Scores.

APPENDIX D

Summary of Means for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 3

Listening

Group	Sex	Y (1,0) 0 Weeks	Y (1,9) 9 Weeks	Y (1,18) 18 Weeks	Y (1,27) 27 Weeks
1	Boys	11.26	38.16	26.64	24.16
	Girls	10.25	35.91	28.13	24.05
2	Boys	3.60	42.00	32.31	33.42
	Girls	2.41	43.37	33.86	33.94
3	Boys	8.82	30.94	19.42	18.43
	Girls	8.54	32.66	21.91	20.28
4	Boys	7.89	39.06	26.53	25.39
	Girls	7.40	38.20	28.46	25.87
5	Boys	4.00	35.66	22.50	24.09
	Girls	10.11	36.47	26.86	26.95
6	Boys	10.09	31.00	21.63	23.83
	Girls	9.40	31.12	24.50	22.07
7	Boys	9.75	36.75	26.50	22.00
	Girls	8.76	39.06	31.20	27.33
8	Boys	14.00	39.38	25.07	25.82
	Girls	12.14	33.25	22.62	20.58
Total	Boys	8.59	36.71	25.30	24.77
	Girls	8.54	37.00	27.84	25.62
Total Boys and Girls		8.57	36.85	26.56	25.19

APPENDIX D

Summary of Means for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 3
(cont'd)

Reading

Group	Sex	Y (2,9) 9 Weeks	Y (2,18) 18 Weeks	Y (2,27) 27 Weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys	33.23	54.14	23.44	1.00	5,33
	Girls	32.30	54.06	26.55	1.00	5.00
2	Boys	40.05	62.05	33.47	1.00	2.31
	Girls	41.64	64.00	34.35	1.00	2.23
3	Boys	28.79	39.00	17.80	2.00	5.21
	Girls	30.16	43.41	19.46	2.00	4.53
4	Boys	33.17	50.53	25.18	2.00	4.35
	Girls	33.31	53.00	27.75	2.00	4.00
5	Boys	29.89	45.33	21.33	2.00	5.11
	Girls	33.76	50.81	26.50	2.00	5.25
6	Boys	28.27	46.09	21.18	3.00	5.56
	Girls	27.87	48.62	20.07	3.00	6.00
7	Boys	32.50	52.50	21.33	3.00	6.20
	Girls	38.66	58.62	28.14	3.00	5.00
8	Boys	32.77	48.46	23.94	1.00	6.00
	Girls	29.28	44.50	19.36	1.00	6.88
Total	Boys	32.69	49.92	23.79	1.79	4.80
	Girls	34.11	53.27	26.00	1.89	4.67
Total Boys and Girls		33.39	51.60	24.90	1.84	4.74

*** San Diego Test Performance Scores.**

APPENDIX D

Summary of Means for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 4 **Listening**

Group	Sex	Y (1,0) 0 Weeks	Y (1,9) 9 Weeks	Y (1,18) 18 Weeks	Y (1,27) 27 Weeks
1	Boys	6.65	35.13	22.11	24.85
	Girls	7.39	37.78	25.79	27.37
2	Boys	6.43	32.55	20.47	20.00
	Girls	6.13	35.93	24.56	22.29
3	Boys	4.84	25.37	17.00	25.36
	Girls	6.63	28.53	24.23	29.46
4	Boys	3.33	37.28	24.71	24.50
	Girls	3.85	37.70	22.88	27.12
5	Boys	8.75	37.25	25.25	27.31
	Girls	10.80	40.83	29.50	29.66
6	Boys	5.50	37.86	20.00	24.93
	Girls	4.30	39.90	25.09	26.35
7	Boys	3.87	39.00	26.00	25.87
	Girls	7.00	41.76	30.00	29.71
8	Boys	13.83	36.61	21.38	26.00
	Girls	13.00	39.80	24.30	27.00
Total	Boys	6.93	34.90	21.94	24.74
	Girls	7.05	37.59	25.65	27.20
Total Boys and Girls		6.98	36.13	23.63	25.85

APPENDIX D

Summary of Means for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 4
(cont'd)

Reading

Group	Sex	Y (2,9) 9 Weeks	Y (2,18) 18 Weeks	Y (2,27) 27 Weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys	30.60	42.26	20.34	2.00	5.34
	Girls	33.92	52.55	24.54	2.00	4.94
2.	Boys	29.27	43.52	17.25	3.00	5.44
	Girls	31.00	49.43	22.00	3.00	4.81
3	Boys	22.68	43.06	16.00	3.00	5.44
	Girls	29.53	52.69	21.00	3.00	4.93
4	Boys	32.85	51.71	23.28	3.00	3.61
	Girls	32.93	49.82	24.18	3.00	4.66
5	Boys	32.81	52.18	21.00	2.00	5.13
	Girls	35.58	59.33	28.50	2.00	4.18
6	Boys	31.40	44.46	20.18	2.00	5.87
	Girls	37.63	47.63	24.42	2.00	4.50
7	Boys	35.63	50.07	23.33	3.00	5.25
	Girls	36.61	59.30	29.28	3.00	4.64
8	Boys	29.52	45.83	21.35	1.00	6.70
	Girls	34.22	49.70	22.40	1.00	6.90
Total	Boys	30.28	47.09	20.17	2.42	5.38
	Girls	33.72	52.46	24.42	2.50	4.91
Total Boys and Girls		31.83	49.48	22.09	2.46	5.16

* San Diego Test Performance Scores.

APPENDIX D

Summary of Mean Raw Scores
for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 5 **Listening**

Group	Sex	Y (1,0) 0 weeks	Y (1,9) 9 weeks	Y (1,18) 18 weeks	Y (1,27) 27 weeks
1	Boys	8.41	37.66	23.63	26.53
	Girls	9.98	38.14	25.57	27.61
2	Boys	7.61	36.08	26.25	26.00
	Girls	9.58	37.93	27.43	26.36
3	Boys	8.46		27.15	
	Girls	9.87		30.88	25.40
4	Boys	8.60	34.73	22.73	24.35
	Girls	9.93	35.50	25.82	24.73
5	Boys	9.90	34.77	20.33	24.07
	Girls	9.38	34.43	21.50	23.73
6	Boys	5.12	29.58	19.91	19.50
	Girls	8.28	35.28	26.71	24.76
7	Boys	6.09	27.36	16.63	15.46
	Girls	7.31	29.26	20.87	22.23
8	Boys	13.54	40.93	31.71	28.87
	Girls	13.91	39.35	32.25	28.42
Total	Boys	8.86	34.75	23.39	23.73
	Girls	9.71	35.68	25.38	25.35
Total Boys and Girls		9.35	35.25	24.46	24.61

APPENDIX D

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 5
(cont'd)

Reading

Group	Sex	Y (2,9) 9 weeks	Y (2,18) 18 weeks	Y (2,27) 27 weeks	Y (3) Tchr.Rat'g	X(1) Ability*
1	Boys	34.81	44.33	24.21	1.00	4.14
	Girls	36.14	51.35	27.55	1.00	3.50
2	Boys	35.08	53.08	22.16	2.00	4.50
	Girls	35.46	55.50	24.36	2.00	4.38
3	Boys		53.23		2.00	4.65
	Girls		60.88		2.00	4.25
4	Boys	31.86	49.00	23.05	1.00	4.75
	Girls	34.94	51.27	24.35	1.00	4.88
5	Boys	31.66	41.00	19.80	2.00	5.50
	Girls	30.94	42.70	20.15	2.00	6.55
6	Boys	24.08	41.83	14.92	1.00	5.40
	Girls	30.00	50.85	22.00	1.00	4.11
7	Boys	20.90	34.45	9.61	3.00	7.00
	Girls	26.75	42.81	16.93	3.00	6.23
8	Boys	34.85	59.33	29.25	2.00	5.35
	Girls	32.78	60.76	28.78	2.00	5.58
Total	Boys	30.64	47.02	20.75	1.69	5.29
	Girls	32.57	50.36	23.43	1.71	5.26
Total Boys and Girls		31.69	48.82	22.20	1.70	5.27

* San Diego Test Performance Scores.

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 6 **Listening**

Group	Sex	Y (1,0) 0 weeks	Y (1,9) 9 weeks	Y (1,18) 18 weeks	Y (1,27) 27 weeks
1	Boys	8.60			
	Girls	9.30			
2	Boys	8.58	38.07	27.46	28.71
	Girls	9.33	40.89	31.15	30.30
3	Boys	8.80	33.62	23.06	22.37
	Girls	9.43	35.80	26.07	25.17
4	Boys	5.44	29.55	18.70	27.38
	Girls	6.14	41.20	34.50	30.90
5	Boys	7.36	38.00	27.14	26.40
	Girls	6.18	40.29	29.94	28.45
6	Boys	8.13	28.28	19.86	21.35
	Girls	9.63	32.41	23.58	27.08
7	Boys	7.87	39.37	23.00	25.60
	Girls	9.50	38.75	26.62	28.52
8	Boys	11.33	36.87	21.37	20.88
	Girls	10.36	40.43	26.05	25.35
Total	Boys	7.60	33.61	22.39	23.95
	Girls	8.39	38.22	27.20	27.33
Total Boys and	Girls	8.05	36.08	24.92	25.76

APPENDIX D

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 6
(cont'd)

Reading

Group	Sex	Y (2,9) 9 weeks	Y (2,18) 18 weeks	Y (2,27) 27 weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys					
	Girls					
2	Boys	36.84	55.61	31.69	2.00	4.92
	Girls	39.05	62.31	31.90	2.00	4.16
3	Boys	30.00	46.18	22.43	3.00	4.50
	Girls	32.40	50.93	26.75	3.00	4.23
4	Boys	25.10	38.00	19.82	3.00	5.10
	Girls	36.66	52.00	27.00	3.00	4.16
5	Boys	32.85	56.35	27.46	1.00	5.00
	Girls	38.70	59.82	28.85	1.00	4.61
6	Boys	25.06	40.73	21.05	2.00	5.25
	Girls	29.45	48.50	28.16	2.00	4.81
7	Boys	32.00	47.62	23.80	2.00	6.12
	Girls	35.66	54.31	28.76	2.00	5.75
8	Boys	31.25	38.62	19.33	2.00	4.75
	Girls	37.41	52.47	24.17	2.00	5.28
Total	Boys	29.59	44.65	22.71	2.16	5.02
	Girls	35.46	53.76	27.06	2.05	4.77
Total Boys and Girls		32.71	49.52	25.07	2.10	4.88

* San Diego Test Performance Scores.

APPENDIX D

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 7 **Listening**

Group	Sex	Y (1,0) 0 weeks	Y (1,9) 9 weeks	Y (1,18) 18 weeks	Y (1,27) 27 weeks
1	Boys	6.52	34.86	24.80	25.35
	Girls	6.66	40.50	27.35	27.41
2	Boys		28.25	15.71	17.47
	Girls	3.00	31.91	20.76	23.88
3	Boys	4.66	31.64	24.92	23.66
	Girls	4.58	37.87	24.33	26.38
4	Boys	9.71	35.81	21.62	21.11
	Girls	11.08	36.75	22.00	24.40
5	Boys	11.71	36.40	22.43	22.76
	Girls	10.55	34.33	21.77	22.90
6	Boys	10.00	36.08	22.83	25.83
	Girls	9.76	38.25	22.78	27.22
7	Boys	11.09	33.61	22.61	26.07
	Girls	14.11	38.50	25.62	27.15
8	Boys		39.29	28.94	28.00
	Girls		40.25	31.43	29.00
Total	Boys	8.80	34.32	22.69	23.57
	Girls	8.72	37.54	24.83	26.22
Total Boys and Girls		8.76	35.77	23.65	24.88

APPENDIX D

Summary of Mean Raw Scores
for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 7
(cont'd)

Reading

Group	Sex	Y (2,9) 9 weeks	Y (2,18) 18 weeks	Y (2, 27) 27 weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys	32.60	49.80	23.35	1.00	4.47
	Girls	39.64	56.00	28.00	1.00	3.87
2	Boys	25.15	30.81	17.42	2.00	4.63
	Girls	30.38	41.23	22.77	2.00	4.35
3	Boys	32.07	47.64	24.06	2.00	4.21
	Girls	34.06	50.37	22.77	2.00	4.23
4	Boys	31.31	46.06	20.22	1.00	4.58
	Girls	31.77	46.75	23.20	1.00	4.75
5	Boys	32.75	44.37	18.88	1.00	5.11
	Girls	32.66	44.33	18.70	1.00	5.10
6	Boys	33.08	49.25	21.66		5.14
	Girls	33.50	48.57	22.05	2.00	6.35
7	Boys	32.36	49.30	25.21	2.00	6.90
	Girls	34.87	52.12	27.00	2.00	6.80
8	Boys	37.33	57.35	26.38	1.00	3.44
	Girls	38.86	61.18	30.16	1.00	3.33
Total	Boys	31.81	46.08	21.97	1.51	4.68
	Girls	34.75	50.60	24.58	1.54	4.68
Total Boys and Girls		33.15	48.12	23.24	1.52	4.68

* San Diego Test Performance Scores.

APPENDIX D

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 8 — Listening

Group	Sex	Y (1,0) 0 weeks	Y (1,9) 9 weeks	Y (1,18) 18 weeks	Y (1,27) 27 weeks
1	Boys	5.21	37.20	28.07	26.36
	Girls	6.66	40.11	27.62	27.60
2	Boys	8.81	36.38	24.46	24.43
	Girls	9.94	41.10	29.73	29.38
3	Boys	7.64	36.23	25.53	23.06
	Girls	6.33	39.15	28.85	27.20
4	Boys	11.26	37.66	23.93	27.47
	Girls	10.43	38.00	24.00	27.11
5	Boys	11.66	36.33	22.33	28.55
	Girls	12.20	30.50	16.00	24.83
6	Boys	10.00	33.72	20.50	22.50
	Girls	13.33	38.40	25.30	24.15
7	Boys	7.18	39.40	26.06	27.73
	Girls	3.58	38.11	23.72	26.33
8	Boys	13.30	41.30	26.33	26.14
	Girls	10.61	38.60	30.11	28.84
Total	Boys	9.31	37.31	24.87	25.68
	Girls	8.88	38.79	26.49	27.12
Total Boys and Girls		9.10	38.03	25.66	26.38

APPENDIX D

Summary of Mean Raw Scores
for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 8
(cont'd)

		Reading				
Group	Sex	Y (2,9) 9 weeks	Y (2,18) 18 weeks	Y (2,27) 27 weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys	33.26	55.46	22.83	1.00	4.88
	Girls	35.55	53.22	28.10	1.00	3.50
2	Boys	35.38	50.69	24.81	3.00	3.62
	Girls	37.63	59.52	30.55	3.00	3.36
3	Boys	31.23	51.60	21.93	2.00	3.06
	Girls	37.38	56.71	27.60	2.00	3.53
4	Boys	33.13	50.73	27.41	1.00	4.75
	Girls	35.07	51.46	25.94	1.00	4.37
5	Boys	32.33	46.22	27.88	3.00	4.77
	Girls	29.00	33.00	22.00	3.00	6.66
6	Boys	29.27	43.54	17.92	3.00	6.63
	Girls	31.90	49.70	23.00	3.00	5.22
7	Boys	34.86	53.66	20.26	1.00	5.53
	Girls	35.23	49.77	19.52	1.00	5.94
8	Boys	33.80	51.50	24.07	3.00	5.84
	Girls	34.54	52.72	25.46	3.00	5.69
Total	Boys	33.03	50.88	23.25	1.95	4.80
	Girls	35.32	52.83	25.29	1.99	4.65
Total Boys and Girls		34.14	51.83	24.25	1.97	4.73

* San Diego Test Performance Scores.

APPENDIX D

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 9 **Listening**

Group	Sex	Y (1,0) 0 weeks	Y (1,9) 9 weeks	Y (1,18) 18 weeks	Y (1,27) 27 weeks
1	Boys	6.14	37.80	30.00	27.56
	Girls	7.80	38.66	27.25	28.31
2	Boys	6.31	34.86	25.43	25.20
	Girls	7.58	37.46	28.26	26.82
3	Boys	8.66	38.87	29.80	29.15
	Girls	8.93	39.21	29.50	28.00
4	Boys	4.60	33.95	22.35	24.23
	Girls	4.90	34.83	24.58	24.15
5	Boys	8.00	36.33	22.75	23.00
	Girls		38.26	27.66	24.27
6	Boys		37.33	23.08	24.07
	Girls		34.70	24.90	23.66
7	Boys	5.75	37.60	21.60	22.26
	Girls	4.55	40.45	26.25	24.28
8	Boys	8.18	31.20	17.66	17.61
	Girls	9.90	30.44	16.60	19.18
Total	Boys	6.54	36.09	24.24	24.36
	Girls	7.48	37.07	26.04	25.15
Total Boys and Girls		6.96	36.54	25.07	24.72

APPENDIX D

Summary of Mean Raw Scores
for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Method 9
(cont'd)

Reading

Group	Sex	Y (2,9) 9 weeks	Y (2,18) 18 weeks	Y (2,27) 27 weeks	Y (3) Tchr.Rat'g	X (1) Ability*
1	Boys	35.66	58.93	27.73		3.87
	Girls	37.50	55.58	28.00	2.00	3.43
2	Boys	31.00	52.06	21.35	3.00	4.37
	Girls	35.33	55.33	26.00	3.00	4.26
3	Boys	36.46	54.56	30.88	2.00	3.00
	Girls	35.35	59.50	29.14	2.00	2.53
4	Boys	32.55	44.60	20.38	1.00	4.42
	Girls	31.66	46.58	20.07	1.00	4.46
5	Boys	29.83	46.25	17.66	3.00	5.00
	Girls	32.73	55.73	22.72	3.00	4.33
6	Boys	31.41	47.91	21.21	1.00	5.38
	Girls	31.40	50.80	22.41	1.00	5.27
7	Boys	33.66	43.53	18.52	2.00	4.89
	Girls	37.81	53.41	23.21	2.00	3.61
8	Boys	24.58	35.41	15.53	2.00	6.30
	Girls	25.70	35.63	18.25	2.00	6.27
Total	Boys	32.17	48.22	21.62	2.06	4.55
	Girls	33.66	52.13	23.92	2.06	4.16
Total Boys and Girls		32.85	50.02	22.66	2.06	4.37

* San Diego Test Performance Scores.

APPENDIX D

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Total Of All 9 Methods **Listening**

	Y (1,0) 0 weeks	Y (1,9) 9 weeks	Y (1,18) 18 weeks	Y (1,27) 27 weeks
Boys	7.88	35.62	23.88	24.50
Girls	8.18	37.51	26.29	26.30
Total Boys and Girls	8.05	36.54	25.04	25.38
	N = 1622	N = 1902	N = 1914	N = 2259

APPENDIX D

Summary of Mean Raw Scores for Each Group Within a Method

COOPERATIVE SPANISH STUDY

Total of All 9 Methods (cont'd)	Y (2,9) 9 weeks	Y (2,18) 18 weeks	Y (2,27) 27 weeks	Y (3) Tchr.Rat'g	X (1) Ability*
		Reading			
Boys	32.00	42.19	22.14	1.89	5.00
Girls	34.58	52.39	24.87	1.89	4.65
Total Boys and Girls	33.26	50.24	23.48	1.89	4.83
	N = 1894	N = 1947	N = 2250	N = 2262	N = 1958

* San Diego Test Performance Scores.